

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2005, 19:46:44 ; Search time 21.6 Seconds  
(without alignments)  
27.633 Million cell updates/sec

Title: US-09-870-216C-3

Perfect score: 44

Sequence: 1 FLQILMEPV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	77.3	318	US-09-583-110-2976	Sequence 2976, Ap
2	33	75.0	305	US-09-710-279-1264	Sequence 1264, Ap
3	33	75.0	324	US-09-134-001C-5525	Sequence 5525, Ap
4	33	75.0	368	US-09-413-574-4	Sequence 4, Appl
5	33	75.0	409	US-09-538-092-1214	Sequence 1214, Ap
6	33	75.0	2802	US-09-542-331-1	Sequence 1, Appl
7	33	75.0	2802	US-09-510-791-1	Sequence 1, Appl
8	32	72.7	107	US-09-328-352-5811	Sequence 5811, Ap
9	32	72.7	195	US-09-134-000C-4338	Sequence 4338, Ap
10	32	72.7	323	US-09-107-532A-5460	Sequence 5460, Ap
11	32	72.7	406	US-09-248-796A-17915	Sequence 17915, A
12	32	72.7	434	US-09-005-286B-2	Sequence 2, Appl
13	32	72.7	642	US-08-706-935-3	Sequence 3, Appl
14	31	70.5	131	US-09-270-767-37595	Sequence 37595, A
15	31	70.5	131	US-09-270-767-52812	Sequence 52812, A
16	31	70.5	375	US-09-252-149B-35	Sequence 35, Appl
17	31	70.5	375	US-09-451-501-27	Sequence 27, Appl
18	31	70.5	375	US-09-686-344-27	Sequence 27, Appl
19	31	70.5	375	US-09-626-896-18	Sequence 18, Appl
20	31	70.5	375	US-09-485-046-14	Sequence 14, Appl
21	31	70.5	1000	US-09-193-562D-30	Sequence 30, Appl
22	31	70.5	1000	US-10-055-412B-10	Sequence 30, Appl
23	31	70.5	1253	US-08-252-966B-12	Sequence 12, Appl
24	31	70.5	1261	US-08-252-966B-18	Sequence 18, Appl
25	30	68.2	103	US-09-302-626B-38	Sequence 38, Appl
26	30	68.2	116	US-09-252-991A-29715	Sequence 29715, A
27	30	68.2	327	US-09-252-991A-33067	Sequence 33067, A

28	30	68.2	475	4	US-09-248-796A-18640	Sequence 18640, A
29	30	68.2	515	4	US-09-434-840-63	Sequence 63, Appl
30	30	68.2	526	4	US-09-434-840-2	Sequence 2, Appl
31	30	68.2	540	4	US-09-252-991A-18305	Sequence 18305, A
32	30	68.2	541	4	US-09-434-840-55	Sequence 55, Appl
33	30	68.2	728	3	US-08-915-337-2	Sequence 2, Appl
34	30	68.2	1332	4	US-09-252-991A-25772	Sequence 25772, A
35	30	68.2	1501	4	US-09-134-001C-3945	Sequence 3945, Ap
36	30	68.2	1529	3	US-09-641-803-2	Sequence 2, Appl
37	29	65.9	17	4	US-09-046-894-38	Sequence 38, Appl
38	29	65.9	44	1	US-08-480-784-83	Sequence 83, Appl
39	29	65.9	45	1	US-08-483-553-83	Sequence 83, Appl
40	29	65.9	45	1	US-08-487-002-83	Sequence 83, Appl
41	29	65.9	45	1	US-08-483-554B-83	Sequence 83, Appl
42	29	65.9	45	1	US-08-488-011B-83	Sequence 83, Appl
43	29	65.9	45	1	US-08-850-727-83	Sequence 83, Appl
44	29	65.9	45	5	PCT-US95-10202-83	Sequence 83, Appl
45	29	65.9	45	5		

#### ALIGNMENTS

```
RESULT 1
US-09-583-110-2976
; Sequence 2976, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2976
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2976

Query Match      77.3%  Score 34;  DB 4;  Length 318;
Best Local Similarity 75.0%  Pred. No. 37;
Matches 6;  Conservative 2;  Mismatches 0;  Indels 0;  Gaps 0;

QY      1 FLQILMEPV 8
Db      152 FLQILMEPV 159

RESULT 2
US-09-710-279-1264
; Sequence 1264, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STREPTOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/09710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1264
; LENGTH: 305
; TYPE: PRT
```

ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: amino acid sequence  
US-09-710-279-1264

Query Match  
Best Local Similarity 75.0%; Score 33; DB 4; Length 305;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMNEP 8  
Db 138 FLOLMNEP 145

RESULT 3  
US-09-134-001C-5525  
Sequence 5525, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 5525  
LENGTH: 324  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5525

Query Match  
Best Local Similarity 75.0%; Score 33; DB 3; Length 324;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMNEP 8  
Db 157 FLOLMNEP 164

RESULT 4  
US-09-413-574-4  
Sequence 4, Application US/09413574  
Patent No. 6235972  
GENERAL INFORMATION:  
APPLICANT: Mahajan, Pramod B.  
TITLE OF INVENTION: Maize Rad23 Genes and Uses Thereof  
FILE REFERENCE: 0964  
CURRENT APPLICATION NUMBER: US/09/413,574  
CURRENT FILING DATE: 1999-10-06  
EARLIER APPLICATION NUMBER: 60/109,728  
EARLIER FILING DATE: 1998-11-23  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 368  
TYPE: PRT  
ORGANISM: Zea mays  
US-09-413-574-4

Query Match  
Best Local Similarity 75.0%; Score 33; DB 3; Length 368;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMNEP 8

Db 291 FLOLMNEP 298

RESULT 5  
US-09-538-092-1214  
Sequence 1214, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Giot, Loic  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CuraPatSeqFormatter Version 0.9  
SEQ ID NO 1214  
LENGTH: 409  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)..(0)  
OTHER INFORMATION: Polypeptide Accession Number P54727  
US-09-538-092-1214

Query Match  
Best Local Similarity 75.0%; Score 33; DB 4; Length 409;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMNEP 9  
Db 324 FLOLMNEP 332

RESULT 6  
US-09-542-331-1  
Sequence 1, Application US/09542331  
Patent No. 6261761  
GENERAL INFORMATION:  
APPLICANT: Zhong, Yi  
APPLICANT: Guo, Hui-Fu  
TITLE OF INVENTION: NF1 Protein and Its Role in Activation  
TITLE OF INVENTION: of Adenyllyl Cyclase by PACAP38-Like Neuropeptides  
FILE REFERENCE: 1314.1047003  
CURRENT APPLICATION NUMBER: US/09/542,331  
CURRENT FILING DATE: 2000-04-04  
EARLIER APPLICATION NUMBER: US 09/046,745  
EARLIER FILING DATE: 1998-03-24  
EARLIER APPLICATION NUMBER: US 60/041,469  
EARLIER FILING DATE: 1997-03-24  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 2802  
TYPE: PRT  
ORGANISM: Drosophila  
US-09-542-331-1

Query Match  
Best Local Similarity 75.0%; Score 33; DB 3; Length 2802;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMNEP 9  
Db 1340 FLOLMNEP 1348

RESULT 7

```
US-09-510-791-1
; Sequence 1, Application US/09510791
; Patent No. 6365126
; GENERAL INFORMATION:
; APPLICANT: Zhong, Yi
; APPLICANT: Guo, Hui-Fu
; APPLICANT: Tong, JiaYuan
; TITLE OF INVENTION: Improvement of Learning and Short Term
; FILE REFERENCE: Memory Defects with Neurofibromatosis 1 (NF1) Expression
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: US/09/510,791
; PRIOR FILING DATE: 1998-03-24
; PRIOR APPLICATION NUMBER: US 09/046,745
; PRIOR FILING DATE: 1997-03-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2802
; TYPE: PRT
; ORGANISM: Drosophila
US-09-510-791-1

Query Match
Best Local Similarity 75.0%; Score 33; DB 3; Length 2802;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FLOLMEPV 9
Db 1340 YLOLMEPV 1348

RESULT 8
US-09-328-352-5811
; Sequence 5811, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5811
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5811

Query Match
Best Local Similarity 72.7%; Score 32; DB 4; Length 107;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLOLMEPV 9
Db 2 YLEKLMRPV 10

RESULT 9
US-09-134-000C-4338
; Sequence 4338, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
```

```
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent version 3.1
; SEQ ID NO 4338
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4338

Query Match
Best Local Similarity 72.7%; Score 32; DB 4; Length 195;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FLOLMEPV 8
Db 27 FLOVIMDP 34

RESULT 10
US-09-107-532A-5460
; Sequence 5460, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arimello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5460:
; LENGTH: 323 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...323
; SEQUENCE DESCRIPTION: SEQ ID NO: 5460:
US-09-107-532A-5460

Query Match
Best Local Similarity 72.7%; Score 32; DB 4; Length 323;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FLOLMEPV 8
```

Db 155 FLOVIMDP 162

RESULT 11  
US-09-248-796A-17915  
; Sequence 17915, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; PRIOR FILING DATE: 1999-02-12  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 17915  
; LENGTH: 406  
; TYPE: PRT  
; ORGANISM: Candida albicans  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (403)  
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unk  
US-09-248-796A-17915

Query Match 72.7%; Score 32; DB 4; Length 406;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 38 FLOLSHPI 46

RESULT 12  
US-09-005-286B-2  
; Sequence 2, Application US/09005286B  
; Patent No. 6756491  
; GENERAL INFORMATION:  
; APPLICANT: EVANS, RONALD M.  
; APPLICANT: BLUMBERG, BRUCE  
; TITLE OF INVENTION: NOVEL STEROID-ACTIVATED NUCLEAR RECEPTORS AND USES  
; FILE REFERENCE: 088802-5201  
; CURRENT APPLICATION NUMBER: US/09/005,286B  
; CURRENT FILING DATE: 1998-01-09  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 434  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-005-286B-2

Query Match 72.7%; Score 32; DB 4; Length 434;  
Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 315 FLOLMEPV 323

RESULT 13  
US-08-706-936-3  
; Sequence 3, Application US/08706936  
; Patent No. 5792851  
; GENERAL INFORMATION:

APPLICANT: VICTOR L. SCHUSTER AND RUN LU  
; TITLE OF INVENTION: HUMAN PROSTAGLANDIN TRANSPORTER  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN  
; STREET: 90 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/706,936  
; FILING DATE: SEPTEMBER 3, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CRAIG J. ARNOLD  
; REGISTRATION NUMBER: 34,287  
; REFERENCE/DOCKET NUMBER: 96700/406  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 697-5995  
; TELEFAX: (212) 286-0854 OR 286-0082  
; TELEX: TWX 710-581-4766  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 642  
; TYPE: AMINO ACID  
; TOPOLOGY: UNKNOWN  
; MOLECULE TYPE: PROTEIN  
; DESCRIPTION: YES  
; HYPOTHEICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: RAT  
; INDIVIDUAL ISOLATE: PROSTAGLANDIN TRANSPORTER  
US-08-706-936-3,

Query Match 72.7%; Score 32; DB 1; Length 642;  
Best Local Similarity 66.7%; Pred. No. 2e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 317 FLOLMEPV 325

RESULT 14  
US-09-270-767-37595  
; Sequence 37595, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 37595  
; LENGTH: 131  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-37595

Query Match 70.5%; Score 31; DB 4; Length 131;  
Best Local Similarity 44.4%; Pred. No. 60;  
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9



Db 90 FVQILVQPI 98

RESULT 15  
US-09-270-767-52812  
; Sequence 52812, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; NUMBER OF SEQ ID NOS: 1999-03-17  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 52812  
; LENGTH: 131  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-52812

Query Match 70.5%; Score 31; DB 4; Length 131;  
Best Local Similarity 44.4%; Pred. No. 60;  
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQILMEPV 9  
Db 90 FVQILVQPI 98

RESULT 16  
US-09-252-149B-35  
; Sequence 35, Application US/09252149B  
; Patent No. 6369201  
; GENERAL INFORMATION:  
; APPLICANT: Barker, Christopher A.  
; TITLE OF INVENTION: IMMUNOLOGICAL METHODS TO MODULAR MYOSTATIN IN  
; TITLE OF INVENTION: VERTEBRATE SUBJECTS  
; FILE REFERENCE: 9001-0042  
; CURRENT APPLICATION NUMBER: US/09/252,149B  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: 60/075,213  
; PRIOR FILING DATE: 1998-02-19  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 35  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Melospiza gallopavo  
US-09-252-149B-35

Query Match 70.5%; Score 31; DB 3; Length 375;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVQILMEPV 9  
Db 13 FVQILVQPI 21

RESULT 17  
US-09-451-501-27  
; Sequence 27, Application US/09451501  
; Patent No. 6468535  
; GENERAL INFORMATION:  
; APPLICANT: Se-jin Lee et al.  
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-8  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:

ADDRESSER: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/451,501  
FILING DATE: 30-Mar-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/795,071  
FILING DATE: <Unknown>  
APPLICATION NUMBER: PCT/US94/03019  
FILING DATE: 18-March-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lisa A. Haile, Ph.D.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/105001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
IMMEDIATE SOURCE:  
CLONE: Turkey GDF-8  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..376  
OTHER INFORMATION:

SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-09-451-501-27

Query Match 70.5%; Score 31; DB 4; Length 375;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVQILMEPV 9  
Db 13 FVQILVQPI 21

RESULT 18  
US-09-686-344-27  
; Sequence 27, Application US/09686344  
; Patent No. 6607884  
; GENERAL INFORMATION:  
; APPLICANT: Lee, Se-jin  
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-8  
; FILE REFERENCE: 07265/144001  
; CURRENT APPLICATION NUMBER: US/09/686,344  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: 08/862,445  
; PRIOR FILING DATE: 1997-05-23  
; PRIOR APPLICATION NUMBER: 08/847,910  
; PRIOR FILING DATE: 1997-04-28  
; PRIOR APPLICATION NUMBER: 08/795,071  
; PRIOR FILING DATE: 1997-02-05  
; PRIOR APPLICATION NUMBER: 08/525,596  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: PCT/US94/03019  
; PRIOR FILING DATE: 1994-03-18  
; PRIOR APPLICATION NUMBER: 08/033,923

```
;; PRIOR FILING DATE: 1993-03-19
;; NUMBER OF SEQ ID NOS: 51
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 27
;; LENGTH: 375
;; TYPE: PRT
;; ORGANISM: Meleagris gallopavo
US-09-686-344-27

Query Match
Best Local Similarity 70.5%; Score 31; DB 4; Length 375;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQILMEPV 9
DB 13 FMOILVHPV 21

RESULT 19
US-09-626-896-18
;; Sequence 18, Application US/09626896
;; Patent No. 6656475
;; GENERAL INFORMATION:
;; APPLICANT: Lee, Se-jin
;; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR RECEPTORS,
;; FILE REFERENCE: JH01470-2
;; CURRENT APPLICATION NUMBER: US/09/626,896
;; PRIOR FILING DATE: 2000-07-27
;; PRIOR APPLICATION NUMBER: 09/485,046
;; PRIOR FILING DATE: 2000-01-31
;; PRIOR APPLICATION NUMBER: PCT/US98/15598
;; PRIOR FILING DATE: 1998-07-28
;; PRIOR APPLICATION NUMBER: 60/054,461
;; PRIOR FILING DATE: 1997-08-01
;; NUMBER OF SEQ ID NOS: 29
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 18
;; LENGTH: 375
;; TYPE: PRT
;; ORGANISM: Meleagris gallopavo
US-09-626-896-18

Query Match
Best Local Similarity 70.5%; Score 31; DB 4; Length 375;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQILMEPV 9
DB 13 FMOILVHPV 21

RESULT 20
US-09-485-046-14
;; Sequence 14, Application US/09485046
;; Patent No. 6696260
;; GENERAL INFORMATION:
;; APPLICANT: The Johns Hopkins University School of Medicine
;; APPLICANT: Lee, Se-jin
;; APPLICANT: McPherron, Alexandra
;; TITLE OF INVENTION: METHODS TO IDENTIFY GROWTH DIFFERENTIATION FACTOR (GDF) RECEPTORS
;; FILE REFERENCE: JH01470-1
;; CURRENT APPLICATION NUMBER: US/09/485,046
;; PRIOR FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: PCT/US98/15598
;; PRIOR FILING DATE: 1998-07-28
;; PRIOR APPLICATION NUMBER: US 06/054,461
;; PRIOR FILING DATE: 1997-08-01
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 14
;; LENGTH: 375
```

```
;; TYPE: PRT
;; ORGANISM: Meleagris gallopavo
US-09-485-046-14

Query Match
Best Local Similarity 70.5%; Score 31; DB 4; Length 375;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQILMEPV 9
DB 13 FMOILVHPV 21

RESULT 21
US-09-193-562D-30
;; Sequence 30, Application US/09193562D
;; Patent No. 6309857
;; GENERAL INFORMATION:
;; APPLICANT: Pauli, Benedicht U.
;; TITLE OF INVENTION: Nucleotide Sequences Encoding Mammalian Calcium
;; FILE REFERENCE: 18617.0052
;; CURRENT APPLICATION NUMBER: US/09/193,562D
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: US/60/065,922
;; PRIOR FILING DATE: 1997-11-17
;; NUMBER OF SEQ ID NOS: 47
;; SEQ ID NO 30
;; LENGTH: 1000
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-193-562D-30

Query Match
Best Local Similarity 70.5%; Score 31; DB 3; Length 1000;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLQILMEPV 9
DB 10 FLSLILSPV 18

RESULT 22
US-10-055-412B-30
;; Sequence 30, Application US/10055412B
;; Patent No. 6692939
;; GENERAL INFORMATION:
;; APPLICANT: Pauli, Benedicht U.
;; TITLE OF INVENTION: Nucleotide Sequences Encoding Mammalian Calcium
;; FILE REFERENCE: 18617.0058
;; CURRENT APPLICATION NUMBER: US/10/055,412B
;; PRIOR FILING DATE: 2001-10-29
;; PRIOR APPLICATION NUMBER: US/09/193,562
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: US/60/065,922
;; PRIOR FILING DATE: 1997-11-17
;; NUMBER OF SEQ ID NOS: 47
;; SEQ ID NO 30
;; LENGTH: 1000
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-055-412B-30

Query Match
Best Local Similarity 70.5%; Score 31; DB 4; Length 1000;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLQILMEPV 9
DB 10 FLSLILSPV 18
```

RESULT 23  
US-08-252-966B-12  
Sequence 12, Application US/08252966B  
Patent No. 5624818  
GENERAL INFORMATION:  
APPLICANT: Eisenman, Robert N.  
APPLICANT: Hurlin, Peter J.  
APPLICANT: Ayer, Donald E.  
TITLE OF INVENTION: Regulatory Proteins that Dimerize with  
TITLE OF INVENTION: Mad or Max  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Christensen, O'Connor, Johnson, and KindnessPLLC  
STREET: 1420 Fifth Ave., Suite 2800  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98101-2347  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,966B  
FILING DATE: 01-JUN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Shelton, Dennis K.  
REGISTRATION NUMBER: 26,997  
REFERENCE/DOCKET NUMBER: FHCRI7694  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 682-8100  
TELEFAX: (206) 224-0779  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: translation of msina cDNA; see Figure 23  
HYPOTHEICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
US-08-252-966B-12

Query Match 70.5%; Score 31; DB 1; Length 1253;  
Best Local Similarity 66.7%; Pred. No. 6.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLLMEPV 9  
Db 1233 FLOLLXSPV 1241

RESULT 24  
US-08-252-966B-18  
Sequence 18, Application US/08252966B  
Patent No. 5624818  
GENERAL INFORMATION:  
APPLICANT: Eisenman, Robert N.  
APPLICANT: Hurlin, Peter J.  
APPLICANT: Ayer, Donald E.  
TITLE OF INVENTION: Regulatory Proteins that Dimerize with  
TITLE OF INVENTION: Mad or Max  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Christensen, O'Connor, Johnson, and KindnessPLLC  
STREET: 1420 Fifth Ave., Suite 2800  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA

ZIP: 98101-2347  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,966B  
FILING DATE: 01-JUN-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Shelton, Dennis K.  
REGISTRATION NUMBER: 26,997  
REFERENCE/DOCKET NUMBER: FHCRI7694  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 682-8100  
TELEFAX: (206) 224-0779  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1261 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: translation of msina3 cDNA; see Figure 29A, B, C, D  
HYPOTHEICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
US-08-252-966B-18

Query Match 70.5%; Score 31; DB 1; Length 1261;  
Best Local Similarity 66.7%; Pred. No. 6.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLLMEPV 9  
Db 1242 FLOLLXSPV 1250

RESULT 25  
US-09-302-626B-38  
Sequence 38, Application US/09302626B  
Patent No. 6709660  
GENERAL INFORMATION:  
APPLICANT: Scalato, Enzo  
APPLICANT: Rappuoli, Rino  
APPLICANT: Piza, Mariagrazia  
APPLICANT: Grandi, Guido  
TITLE OF INVENTION: Meningococcal Antigens  
FILE REFERENCE: CHIR0159  
CURRENT APPLICATION NUMBER: US/09/302,626B  
PRIOR FILING DATE: 1999-04-30  
PRIOR APPLICATION NUMBER: PCT/IB99/00103  
NUMBER OF SEQ ID NOS: 195  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 38  
LENGTH: 103  
TYPE: PRT  
ORGANISM: Neisseria meningitidis  
US-09-302-626B-38

Query Match 68.2%; Score 30; DB 4; Length 103;  
Best Local Similarity 66.7%; Pred. No. 75;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLLMEPV 9  
Db 26 FLOLLXSPV 34

RESULT 26

US-09-252-991A-29715  
; Sequence 29715, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 29715  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29715

Query Match  
Best Local Similarity 68.2%; Score 30; DB 4; Length 116;  
Best Local Similarity 75.0%; Pred. No. 85;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LQLEMEPV 9  
Db 91 LQLEMEPI 98

RESULT 27  
US-09-252-991A-33067  
; Sequence 33067, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 33067  
; LENGTH: 327  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-33067

Query Match  
Best Local Similarity 68.2%; Score 30; DB 4; Length 327;  
Best Local Similarity 85.7%; Pred. No. 2.5e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLQLLME 7  
Db 210 FLQLLBL 216

RESULT 28  
US-09-248-796A-18640  
; Sequence 18640, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248.796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 18640  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-18640

Query Match  
Best Local Similarity 68.2%; Score 30; DB 4; Length 475;  
Best Local Similarity 62.5%; Pred. No. 3.7e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 LQLEMEPV 9  
Db 98 LQLEMEPI 105

RESULT 29  
US-09-434-840-63  
; Sequence 63, Application US/09434840  
; Patent No. 6620985  
; GENERAL INFORMATION:  
; APPLICANT: Glazebrook, Jane  
; APPLICANT: Jirag, Dayadevi  
; APPLICANT: Toocle, Tina L  
; APPLICANT: Zhou, Nan  
; APPLICANT: Feys, Bart  
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR  
; FILE REFERENCE: 043503.0009  
; CURRENT APPLICATION NUMBER: US/09/434.840  
; CURRENT FILING DATE: 1999-11-04  
; EARLIER APPLICATION NUMBER: 09/190,733  
; EARLIER FILING DATE: 1998-11-12  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 63  
; LENGTH: 515  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-434-840-63

Query Match  
Best Local Similarity 68.2%; Score 30; DB 4; Length 515;  
Best Local Similarity 66.7%; Pred. No. 4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 FLQLEMEPV 9  
Db 387 FYQLEMEPL 395

RESULT 30  
US-09-434-840-2  
; Sequence 2, Application US/09434840  
; Patent No. 6620985  
; GENERAL INFORMATION:  
; APPLICANT: Glazebrook, Jane  
; APPLICANT: Jirag, Dayadevi  
; APPLICANT: Toocle, Tina L  
; APPLICANT: Zhou, Nan  
; APPLICANT: Feys, Bart  
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR  
; FILE REFERENCE: 043503.0009  
; CURRENT APPLICATION NUMBER: US/09/434.840  
; CURRENT FILING DATE: 1999-11-04  
; EARLIER APPLICATION NUMBER: 09/190,733  
; EARLIER FILING DATE: 1998-11-12  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 526  
; TYPE: PRT

ORGANISM: Arabidopsis thaliana  
US-09-434-840-2

Query Match 68.2%; Score 30; DB 4; Length 526;  
Best Local Similarity 66.7%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 372 FYQLAEPL 380

RESULT 31  
US-09-252-991A-18305  
Sequence 18305, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 18305  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18305

Query Match 68.2%; Score 30; DB 4; Length 540;  
Best Local Similarity 66.7%; Pred. No. 4.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 163 FLPLQGPV 171

RESULT 32  
US-09-434-840-55  
Sequence 55, Application US/09434840  
Patent No. 6620985  
GENERAL INFORMATION:  
APPLICANT: Glazebrook, Jane  
APPLICANT: Jirage, Dayadevi  
APPLICANT: Toocle, Tina L  
APPLICANT: Zhou, Nan  
TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR  
FILE REFERENCE: 043503.0009  
CURRENT APPLICATION NUMBER: US/09/434,840  
CURRENT FILING DATE: 1999-11-04  
EARLIER APPLICATION NUMBER: 09/190,733  
EARLIER FILING DATE: 1998-11-12  
NUMBER OF SEQ ID NOS: 85  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 55  
LENGTH: 541  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-434-840-55

Query Match 68.2%; Score 30; DB 4; Length 541;  
Best Local Similarity 66.7%; Pred. No. 4.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 163 FLPLQGPV 171

Db 387 FYQLAEPL 395

RESULT 33  
US-08-915-337-2  
Sequence 2, Application US/08915337  
Patent No. 6287802  
GENERAL INFORMATION:  
APPLICANT: Deng, Fan & Xia  
TITLE OF INVENTION: EXT2 Gene  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/915,337  
FILING DATE: August 21, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: China - 96121928.9  
FILING DATE: October 21, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: William T. Han  
REGISTRATION NUMBER: 34,344  
REFERENCE/DOCKET NUMBER: ATG50019  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 270-5219  
TELEFAX: (610) 270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 728 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
US-08-915-337-2

Query Match 68.2%; Score 30; DB 3; Length 728;  
Best Local Similarity 85.7%; Pred. No. 5.8e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 QLMMEPV 9  
Db 391 QLMMEPV 397

RESULT 34  
US-09-252-991A-25772  
Sequence 25772, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 25772  
LENGTH: 1332  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25772

Query Match 68.2%; Score 30; DB 4; Length 1332;  
Best Local Similarity 75.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LQ1MEPV 9  
Db 327 LQ1MEPV 334

RESULT 35  
US-09-710-279-2850  
Sequence 2850, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMMERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: P03480US  
CURRENT APPLICATION NUMBER: US/09/710,279  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2850  
LENGTH: 1501  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-2850

Query Match 68.2%; Score 30; DB 4; Length 1501;  
Best Local Similarity 75.0%; Pred. No. 1.2e+03;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLO1MEP 8  
Db 297 FLS1AMEP 304

RESULT 36  
US-09-134-001C-3945  
Sequence 3945, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3945  
LENGTH: 1529  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3945

Query Match 68.2%; Score 30; DB 3; Length 1529;

Best Local Similarity 75.0%; Pred. No. 1.3e+03;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLO1MEP 8  
Db 325 FLS1AMEP 332

RESULT 37  
US-09-641-803-2  
Sequence 2, Application US/09641803  
Patent No. 650798  
GENERAL INFORMATION:  
APPLICANT: STANTON, G. John  
APPLICANT: HUGHES, Thomas K.  
APPLICANT: BOLDOGH, Istvan  
TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND  
TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS  
FILE REFERENCE: 265,00220101  
CURRENT APPLICATION NUMBER: US/09/641,803  
CURRENT FILING DATE: 2000-08-17  
PRIOR APPLICATION NUMBER: 60/149,310  
PRIOR FILING DATE: 1999-08-17  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 17  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-641-803-2

Query Match 65.9%; Score 29; DB 4; Length 17;  
Best Local Similarity 71.4%; Pred. No. 18;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQ1MEP 8  
Db 8 LQ1MEP 14

RESULT 38  
US-09-046-894-38  
Sequence 38, Application US/09046894  
Patent No. 6190857  
GENERAL INFORMATION:  
APPLICANT: Ralph, David  
APPLICANT: An, Gang  
APPLICANT: O'Hara, Mark S.  
APPLICANT: Veltri, Robert  
TITLE OF INVENTION: DIAGNOSIS OF DISEASE STATE USING mRNA  
TITLE OF INVENTION: PROFILES IN PERIPHERAL LEUKOCYTES  
NUMBER OF SEQUENCES: 55  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/046,894  
FILING DATE: Concurrently Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/041,576

FILED DATE: 24-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Nakashima, Richard A.  
REGISTRATION NUMBER: P-42,023  
REFERENCE/DOCKET NUMBER: UROC:014  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 44 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-09-046-894-38

Query Match 65.9%; Score 29; DB 3; Length 44;  
Best Local Similarity 75.0%; Pred. No. 49;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEMPV 9  
|:|:|:|:|  
DB 5 LELKEPV 12

RESULT 39  
US-08-480-784-83  
Sequence 83, Application US/08480784  
Patent No. 5693473  
GENERAL INFORMATION:  
APPLICANT: Skolnick, Mark H.  
APPLICANT: Goldgar, David E.  
APPLICANT: Miki, Yoshio  
APPLICANT: Swenson, Jeff  
APPLICANT: Kamb, Alexander  
APPLICANT: Harshman, Keith D.  
APPLICANT: Shattuck-Eidens, Donna M.  
APPLICANT: Tavtigian, Sean V.  
APPLICANT: Wiseman, Roger W.  
APPLICANT: Futreal, P. Andrew  
TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer  
NUMBER OF SEQUENCES: 85  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, N.W., Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,784  
FILING DATE:  
CLASSIFICATION: 435  
APPLICATION NUMBER: US 08/409,305  
FILING DATE: 24-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/348,824  
FILING DATE: 29-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/308,104  
FILING DATE: 16-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/300,266  
FILING DATE: 02-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/289,221

FILED DATE: 12-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Ihnen, Jeffrey L.  
REGISTRATION NUMBER: 28,957  
REFERENCE/DOCKET NUMBER: 24884-109347  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-962-4810  
TELEFAX: 202-962-8300  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-08-480-784-83

Query Match 65.9%; Score 29; DB 1; Length 45;  
Best Local Similarity 75.0%; Pred. No. 50;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEMPV 9  
|:|:|:|:|  
DB 5 LELKEPV 12

RESULT 40  
US-08-483-553-83  
Sequence 83, Application US/08483553  
Patent No. 5709999  
GENERAL INFORMATION:  
APPLICANT: Skolnick, Mark H.  
APPLICANT: Goldgar, David E.  
APPLICANT: Miki, Yoshio  
APPLICANT: Swenson, Jeff  
APPLICANT: Kamb, Alexander  
APPLICANT: Harshman, Keith D.  
APPLICANT: Shattuck-Eidens, Donna M.  
APPLICANT: Tavtigian, Sean V.  
APPLICANT: Wiseman, Roger W.  
APPLICANT: Futreal, P. Andrew  
TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer  
NUMBER OF SEQUENCES: 85  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Venable, Baetjer, Howard & Civiletti, LLP  
STREET: 1201 New York Avenue, N.W., Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,553  
FILING DATE:  
CLASSIFICATION: 435  
APPLICATION NUMBER: US 08/409,305  
FILING DATE: 24-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/348,824  
FILING DATE: 29-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/308,104  
FILING DATE: 16-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/300,266  
FILING DATE: 02-SEP-1994

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/289,221  
 FILING DATE: 12-AUG-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ihnen, Jeffrey L.  
 REGISTRATION NUMBER: 28,957  
 REFERENCE/DOCKET NUMBER: 24884-109347  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-962-4810  
 TELEFAX: 202-962-8300  
 INFORMATION FOR SEQ ID NO: 83:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 45 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULAR TYPE: peptide  
 HYPOTHEICAL: NO  
 US-08-483-553-83

Query Match 65.9%; Score 29; DB 1; Length 45;  
 Best Local Similarity 75.0%; Pred. No. 50;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LQIMEPV 9  
 DB 5 LELKEPV 12

Search completed: January 12, 2005, 20:17:37  
 Job time : 27.6 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:37:43 ; Search time 201.4 Seconds  
(without alignments)  
16.031 Million cell updates/sec

Title: US-09-870-216C-3

Perfect score: 44

Sequence: 1 FLQJLMEPV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: geneseqp1980a:\*  
2: geneseqp1980a:\*  
3: geneseqp2000a:\*  
4: geneseqp2001a:\*  
5: geneseqp2002a:\*  
6: geneseqp2003a:\*  
7: geneseqp2003b:\*  
8: geneseqp2004a:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	9	5	ABB08361
2	44	100.0	9	7	ABR82213
3	44	100.0	352	5	ABR08366
4	35	79.5	67	4	AAU22301
5	35	79.5	67	7	ADR46269
6	35	79.5	329	6	ABB99785
7	34	77.3	307	5	ABP53033
8	34	77.3	318	5	ABG61495
9	34	77.3	318	5	ABU02362
10	34	77.3	318	8	ADK46461
11	34	77.3	2785	3	AAV57148
12	34	77.3	117	3	ABP34618
13	33	75.0	117	3	ABP34618
14	33	75.0	296	3	AA644343
15	33	75.0	305	4	AA682085
16	33	75.0	318	6	ABW72286
17	33	75.0	324	5	ABR40680
18	33	75.0	345	5	AA644342
19	33	75.0	368	3	AAV71459
20	33	75.0	368	3	AA644341
21	33	75.0	409	2	AAW57000
22	33	75.0	409	2	AAW68186
23	33	75.0	409	6	ABU07460
24	33	75.0	409	8	ADN03716
25	33	75.0	416	5	ABB57171

26	33	75.0	2764	4	ABB66967	ABb66967 Drosophi1
27	33	75.0	2802	4	AAE05485	AAe05485 Drosophi1
28	33	75.0	2802	4	ABB63789	ABb63789 Drosophi1
29	33	75.0	2802	5	ABBO8077	ABb08077 Drosophi1
30	32	72.7	107	6	ADA34524	AdA34524 Acinetoba
31	32	72.7	195	7	ADH86453	ADh86453 Enterococ
32	32	72.7	316	5	AAV42689	AAv42689 H186-preg
33	32	72.7	316	5	AAW50626	AAw50626 H186-preg
34	32	72.7	323	5	ADC95833	AdC95833 E. faeciu
35	32	72.7	329	5	AAU78295	AAu78295 Human pre
36	32	72.7	414	2	AAV42691	AAv42691 Human pre
37	32	72.7	434	2	AAV15931	AAv15931 A human i
38	32	72.7	434	2	AAV21799	AAv21799 Human sce
39	32	72.7	434	4	AAH84417	AAh84417 Amino aci
40	32	72.7	434	5	AAU78294	AAu78294 Human pre
41	32	72.7	434	5	AAW50624	AAw50624 Human pre
42	32	72.7	434	6	AAE31708	AAe31708 Human Stre
43	32	72.7	434	6	ABB99786	ABb99786 Amino aci
44	32	72.7	434	7	ABO23618	ABo23618 Human oip
45	32	72.7	434	7	ABO23620	ABo23620 Human oip

#### ALIGNMENTS

RESULT 1	ABB08361	standard, protein; 9 AA.
ID	ABB08361	standard, protein; 9 AA.
XX		
AC	ABB08361;	
DT	07-MAY-2002 (first entry)	
DE	Synthetic epitope 1 of human cancer antigen eir3.	
XX		
KW	Human; melanoma antigen eukaryotic initiation factor 3; eir3;	
KW	ovarian cancer; MHC; cytosolic; immunomodulator; immune effector cell;	
KW	anti-cancer; vaccine.	
XX		
OS	Homo sapiens.	
FT	Domain	Location/Qualifiers
FT	1	/note= "HLA-2 binding residue"
FT	2	/note= "HLA-2 binding residue"
FT	3..8	/note= "T-cell receptor (TCR) binding domain"
FT	9	/note= "HLA-2 binding residue"
FT	Domain	
XX		
XX	MO200192307-A2.	
XX	06-DEC-2001.	
XX	30-MAY-2001; 2001WO-US017456.	
XX	31-MAY-2000; 2000US-0209391P.	
XX	17-AUG-2000; 2000US-0226258P.	
XX	20-DEC-2000; 2000US-0257088P.	
XX	(GENZ) GENZYME CORP.	
XX	Niclette CA;	
XX	WPI; 2002-139606/18.	
XX	New therapeutic compounds useful against human ovarian cancer, for	
XX	modulating immune response in a subject, and for generating antibodies	
XX	that specifically recognize and bind to these molecules.	
XX	Claim 29; Page 59; 68pp; English.	

CC The invention relates to novel therapeutic compounds, that are designed  
CC to enhance binding to MHC molecules and to enhance immunoregulatory  
CC properties relative to their natural counterparts. The activity of the  
CC compounds of the invention may be described as cytostatic and  
CC immunomodulatory. The compounds are useful against human ovarian cancer,  
CC for modulating immune response in a subject, and for generating  
CC antibodies that specifically recognize and bind to these molecules.  
CC Compositions comprising the compounds are useful as components of anti-  
CC cancer vaccines and to expand immune effector cells that are specific for  
CC cells characterized by expression of antigen eIF3 (melanoma antigen  
CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
CC a detectable agent may be used in diagnostic procedures, such as in the  
CC detection and purification of antibodies, and as immunogens for  
CC production of antibodies. The polynucleotides can be used as primers for  
CC detection of genes or gene transcripts expressed in APC to confirm  
CC transduction of the polynucleotides into host cells. The current sequence  
CC represents synthetic epitope 1 of human cancer antigen eIF3  
SQ Sequence 9 AA;

Query Match 100.0%; Score 44; DB 5; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 1 FLOLMEPV 9  
1 FLOLMEPV 9

RESULT 2  
ABR82213  
ID ABR82213 standard; peptide; 9 AA.

AC ABR82213;  
DT 13-OCT-2003 (first entry)

DE Human antigen eIF3 derived compound 1.

KW Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;  
KM cytosolic; gene therapy; human; antigen.

OS Synthetic.  
OS Homo sapiens.

PN WO2003050543-A1.

PD 19-JUN-2003.

PF 05-DEC-2001; 2001WO-US047997.

PR 05-DEC-2001; 2001WO-US047997.

PA (GENZ ) GENZYME CORP.

PI Nicolette CA;

PR 17-AUG-2000; 2000US-0226258P.

PR 20-DEC-2000; 2000US-0257008P.

PT Aiding in the diagnosis of a neoplastic condition, useful for treating  
PT cancer and related malignancies comprises determining the amount of  
PT expression of an eIF3 protein in a test sample isolated from the cell or  
PT tissue.

PS Claim 12; Page 30; 77pp; English.

CC The invention relates to aiding in the diagnosis of a neoplastic  
CC condition or susceptibility to a neoplastic condition of an animal cell  
CC or tissue. The method involves determining the amount of expression of an  
CC eukaryotic translation initiation factor 3 (eIF3) protein in a test  
CC sample isolated from the cell or tissue, and diagnosing a neoplastic  
CC condition or susceptibility to a neoplastic condition based on the amount

CC of expression of the eIF3 protein. The methods, compounds and kits are  
CC useful in therapeutics, diagnostic and screening methods for human cancer  
CC and related malignancies, e.g. ovarian, breast, lung, colon, prostate,  
CC pancreatic or gastrointestinal cancer, or melanoma. Sequences ABR82213-16  
CC represent compounds derived from the human antigen eIF3  
SQ Sequence 9 AA;

Query Match 100.0%; Score 44; DB 7; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 1 FLOLMEPV 9  
1 FLOLMEPV 9

RESULT 3  
ABB08366  
ID ABB08366 standard; protein; 352 AA.

AC ABB08366;

DT 07-MAY-2002 (first entry)

DE Human cancer antigen eIF3 variant 1 amino acid sequence.

KW Human; melanoma antigen eukaryotic initiation factor 3; eIF3;  
KM ovarian cancer; MHC; cytosolic; immunomodulatory; immune effector cell;  
KM anti-cancer; vaccine.

OS Homo sapiens.

OS Key Location/Qualifiers

FT Misc-difference 242 /note= "wild-type Asn is replaced by Phe"

FT Domain 242 /note= "HLA-2 binding residue"

FT Domain 243 /note= "HLA-2 binding residue"

FT Domain 244,249 /note= "T-cell receptor (TCR) binding domain"

FT Misc-difference 248 /note= "wild-type Asp is replaced by Glu"

FT Misc-difference 249 /note= "wild-type Arg is replaced by Pro"

FT Domain 250 /note= "HLA-2 binding residue"

PN WO200192307-A2.

PD 06-DEC-2001.

PF 30-MAY-2001; 2001WO-US017456.

PR 31-MAY-2000; 2000US-0209391P.

PR 17-AUG-2000; 2000US-0226258P.

PR 20-DEC-2000; 2000US-0257008P.

PA (GENZ ) GENZYME CORP.

PI Nicolette CA;

PR 17-AUG-2000; 2000US-0226258P.

PR 20-DEC-2000; 2000US-0257008P.

PT New therapeutic compounds useful against human ovarian cancer, for  
PT modulating immune response in a subject, and for generating antibodies  
PT that specifically recognize and bind to these molecules.

PS Claim 6; Page; 68pp; English.

CC The invention relates to novel therapeutic compounds, that are designed  
CC to enhance binding to MHC molecules and to enhance immunoregulatory

CC properties relative to their natural counterparts. The activity of the  
CC compounds of the invention may be described as cytostatic and  
CC immunomodulatory. The compounds are useful against human ovarian cancer,  
CC for modulating immune response in a subject, and for generating  
CC antibodies that specifically recognize and bind to these molecules.  
CC Compositions comprising the compounds are useful as components of anti-  
CC cancer vaccines and to expand immune effector cells that are specific for  
CC cells characterised by expression of antigen E1F3 (melanoma antigen  
CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
CC a detectable agent may be used in diagnostic procedures, such as in the  
CC detection and purification of antibodies, and as immunogens for  
CC production of antibodies. The polynucleotides can be used as primers for  
CC detecting genes or gene transcripts expressed in APC to confirm  
CC transduction of the polynucleotides into host cells. The current sequence  
CC represents the human cancer antigen e1f3 variant 1 amino acid sequence.  
CC Note: This sequence is not present in the specification, but may be  
CC created from the sequence of the wild-type human cancer antigen e1f3  
CC sequence given in ABB08360  
CC  
SQ Sequence 352 AA;

Query Match 100.0%; Score 44; DB 5; Length 352;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLIMEPV 9  
DB 242 FLOLIMEPV 250

RESULT 4  
AAU22301  
ID AAU22301 standard; protein; 67 AA.

AC AAU22301;

XX 18-DEC-2001 (first entry)

DE Human cardiovascular system antigen polypeptide SEQ ID No 1075.

XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnereary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
KW anti-fertility.

XX Homo sapiens.

XX WO200155321-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001340.

XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226861P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228287P.  
PR 01-SEP-2000; 2000US-0228343P.  
PR 01-SEP-2000; 2000US-0228344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0233688P.  
PR 14-SEP-2000; 2000US-0233397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-023400P.  
PR 14-SEP-2000; 2000US-0234240P.  
PR 14-SEP-2000; 2000US-0234241P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236377P.  
PR 29-SEP-2000; 2000US-0236377P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237033P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.

PR	20-OCT-2000	2000US-0241825P.
PR	01-NOV-2000	2000US-024617P.
PR	08-NOV-2000	2000US-0246474P.
PR	08-NOV-2000	2000US-0246475P.
PR	08-NOV-2000	2000US-0246476P.
PR	08-NOV-2000	2000US-0246477P.
PR	08-NOV-2000	2000US-0246478P.
PR	08-NOV-2000	2000US-0246523P.
PR	08-NOV-2000	2000US-0246524P.
PR	08-NOV-2000	2000US-0246525P.
PR	08-NOV-2000	2000US-0246526P.
PR	08-NOV-2000	2000US-0246527P.
PR	08-NOV-2000	2000US-0246528P.
PR	08-NOV-2000	2000US-0246532P.
PR	08-NOV-2000	2000US-0246609P.
PR	08-NOV-2000	2000US-0246610P.
PR	08-NOV-2000	2000US-0246611P.
PR	08-NOV-2000	2000US-0246613P.
PR	17-NOV-2000	2000US-0249207P.
PR	17-NOV-2000	2000US-0249208P.
PR	17-NOV-2000	2000US-0249209P.
PR	17-NOV-2000	2000US-0249210P.
PR	17-NOV-2000	2000US-0249211P.
PR	17-NOV-2000	2000US-0249212P.
PR	17-NOV-2000	2000US-0249213P.
PR	17-NOV-2000	2000US-0249214P.
PR	17-NOV-2000	2000US-0249215P.
PR	17-NOV-2000	2000US-0249216P.
PR	17-NOV-2000	2000US-0249217P.
PR	17-NOV-2000	2000US-0249218P.
PR	17-NOV-2000	2000US-0249244P.
PR	17-NOV-2000	2000US-0249245P.
PR	17-NOV-2000	2000US-0249246P.
PR	17-NOV-2000	2000US-0249265P.
PR	17-NOV-2000	2000US-0249297P.
PR	17-NOV-2000	2000US-0249299P.
PR	17-NOV-2000	2000US-0249300P.
PR	01-DEC-2000	2000US-0250160P.
PR	01-DEC-2000	2000US-0250391P.
PR	05-DEC-2000	2000US-0251030P.
PR	05-DEC-2000	2000US-0251988P.
PR	05-DEC-2000	2000US-0256719P.
PR	06-DEC-2000	2000US-0251719P.
PR	08-DEC-2000	2000US-0251856P.
PR	08-DEC-2000	2000US-0251868P.
PR	08-DEC-2000	2000US-0251869P.
PR	08-DEC-2000	2000US-0251989P.
PR	08-DEC-2000	2000US-0251990P.
PR	11-DEC-2000	2000US-0254097P.
PR	05-JAN-2001	2001US-0259678P.

(HUMA-), HUMAN GENOME SCI INC.  
 Rosen CA, Barash SC, Ruben SM;  
 MPI: 2001-451930/48.  
 DR N-PsDB; AAS35575.  
 XX  
 PT New cardiovascular system related polynucleotides and polypeptides,  
 PT useful for diagnosing, treating and/or preventing disorders of the  
 PT cardiovascular system.  
 XX  
 PS Claim 11; SEQ ID NO 1075; 674pp; English.  
 XX  
 CC Sequences AAU21852-AAU22466 represent the cardiovascular system antigen  
 CC polypeptides of the invention. Cardiovascular system antigens and their  
 CC associated polynucleotides are useful in the diagnosis, treatment and  
 CC prevention of various types of disorders in e.g. humans, mice, rabbits,  
 CC goats, horses, cats, dogs, chickens or sheep. A pathological condition  
 CC can be determined by detecting the presence or absence of a mutation in a  
 CC cardiovascular system antigen polynucleotide. The treatable disorders  
 CC include autoimmune diseases such as rheumatoid arthritis,  
 CC hyperproliferative disorders such as neoplasms of the breast or liver,  
 CC

CC cardiovascular disorders such as cardiac arrest, cerebrovascular  
CC disorders such as cerebral ischaemia, nervous system disorders such as  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,  
CC ocular disorders such as corneal infection, endocrine disorders such as  
CC premature labour and infertility, gastrointestinal disorders such as  
CC Crohn's disease, renal disorders such as glomerulonephritis and  
CC respiratory disorders such as asthma and pleurisy. The polypeptides can  
CC also be used to aid wound healing, to prevent skin aging due to sunburn,  
CC to maintain organs before transplantation, to regenerate tissues and in  
CC chemotaxis. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from [ftp.wipo.int/pub/published\\_pat\\_sequences](http://ftp.wipo.int/pub/published_pat_sequences)  
XX

Query Match	79.5%	Score 35;	DB 4;	Length 67;
Best Local Similarity	77.8%	Pred. NO. 15;		
Matches	7;	Conservative	1;	Indels 0; Gaps 0
QY	1	FLQIMEPV	9	
		:		
DB	7	FLQILTPV	15	

RESULT 5  
ADE46269  
ID ADE46269 standard; protein; 67 AA

AC ADE46269

DT 29-JAN-2004 (first entry)

Human cardiovascular system related polypeptide #450

Human; cardiovascular system related polypeptide; cancer; KW

haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;

KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;

cardiovascular disorder; angiogenic disorder; kidney disorder;

**KW** endocrine disorder.

OS Homo sapiens.

US2003059908-A1

PD 27-MAR-2003.

07-MAR-2002; 2002US-00091504.

PR 31-JAN-2000; 2000US-0179065P.

PR 24-FEB-2000; 2000US-0184664P.

16-MAR-2000; 2000US-0189874P. PR

PR 18-APR-2000; 2000US-0198123P.

PR 07-JUN-2000; 2000US-0209467P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217496P

26-JUL-2000; 2000US-0220963P  
PR

PR 14-AUG-2000; 2000US-022451BP

PR 14-AUG-2000; 2000US-0225213P

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

PR	14-AUG-2000	2000US-0225266F
PR	14-AUG-2000	2000US-0225267F
PR	14-AUG-2000	2000US-0225268F
PR	14-AUG-2000	2000US-0225270F
PR	14-AUG-2000	2000US-0225447P
PR	14-AUG-2000	2000US-0225575F
PR	14-AUG-2000	2000US-0225578P
PR	14-AUG-2000	2000US-0225579P
PR	18-AUG-2000	2000US-0226279P
PR	22-AUG-2000	2000US-0226681P
PR	22-AUG-2000	2000US-0226682P
PR	23-AUG-2000	2000US-0227009P
PR	30-AUG-2000	2000US-0228924P
PR	01-SEP-2000	2000US-0229287P
PR	01-SEP-2000	2000US-0229343P
PR	01-SEP-2000	2000US-0229344P
PR	01-SEP-2000	2000US-0229345P
PR	05-SEP-2000	2000US-0229509P
PR	05-SEP-2000	2000US-0229513P
PR	06-SEP-2000	2000US-0230437P
PR	06-SEP-2000	2000US-0230438P
PR	08-SEP-2000	2000US-0231242P
PR	08-SEP-2000	2000US-0231243P
PR	08-SEP-2000	2000US-0231244P
PR	08-SEP-2000	2000US-0231413P
PR	08-SEP-2000	2000US-0231414P
PR	08-SEP-2000	2000US-0232080P
PR	12-SEP-2000	2000US-0231568P
PR	14-SEP-2000	2000US-02323297P
PR	14-SEP-2000	2000US-0232398P
PR	14-SEP-2000	2000US-0232399P
PR	14-SEP-2000	2000US-0232400P
PR	14-SEP-2000	2000US-0232401P
PR	14-SEP-2000	2000US-0233063P
PR	14-SEP-2000	2000US-0233064P
PR	14-SEP-2000	2000US-0233065P
PR	21-SEP-2000	2000US-0234223P
PR	21-SEP-2000	2000US-0234374P
PR	25-SEP-2000	2000US-0234977P
PR	25-SEP-2000	2000US-0234988P
PR	26-SEP-2000	2000US-0235584P
PR	27-SEP-2000	2000US-0235634P
PR	27-SEP-2000	2000US-0235635P
PR	29-SEP-2000	2000US-0236367P
PR	29-SEP-2000	2000US-0236368P
PR	29-SEP-2000	2000US-0236369P
PR	29-SEP-2000	2000US-0236370P
PR	02-OCT-2000	2000US-0236802P
PR	02-OCT-2000	2000US-0237037P
PR	02-OCT-2000	2000US-0237038P
PR	02-OCT-2000	2000US-0237040P
PR	13-OCT-2000	2000US-0239935P
PR	13-OCT-2000	2000US-0239937P
PR	20-OCT-2000	2000US-0240560P
PR	20-OCT-2000	2000US-0241221P
PR	20-OCT-2000	2000US-0241785P
PR	20-OCT-2000	2000US-0241786P
PR	08-NOV-2000	2000US-0246474P
PR	08-NOV-2000	2000US-0246475P
PR	08-NOV-2000	2000US-0246476P
PR	08-NOV-2000	2000US-0246477P
PR	08-NOV-2000	2000US-0246478P
PR	08-NOV-2000	2000US-0246524P

PR	08-NOV-2000	2000US-0246525P.
PR	08-NOV-2000	2000US-0246526P.
PR	08-NOV-2000	2000US-0246527P.
PR	08-NOV-2000	2000US-0246528P.
PR	08-NOV-2000	2000US-0246532P.
PR	08-NOV-2000	2000US-0246609P.
PR	08-NOV-2000	2000US-0246610P.
PR	08-NOV-2000	2000US-0246611P.
PR	08-NOV-2000	2000US-0246613P.
PR	17-NOV-2000	2000US-0249207P.
PR	17-NOV-2000	2000US-0249208P.
PR	17-NOV-2000	2000US-0249209P.
PR	17-NOV-2000	2000US-0249210P.
PR	17-NOV-2000	2000US-0249211P.
PR	17-NOV-2000	2000US-0249212P.
PR	17-NOV-2000	2000US-0249213P.
PR	17-NOV-2000	2000US-0249214P.
PR	17-NOV-2000	2000US-0249215P.
PR	17-NOV-2000	2000US-0249216P.
PR	17-NOV-2000	2000US-0249217P.
PR	17-NOV-2000	2000US-0249218P.
PR	17-NOV-2000	2000US-0249244P.
PR	17-NOV-2000	2000US-0249245P.
PR	17-NOV-2000	2000US-0249264P.
PR	17-NOV-2000	2000US-0249265P.
PR	17-NOV-2000	2000US-0249297P.
PR	17-NOV-2000	2000US-0249299P.
PR	17-NOV-2000	2000US-0249300P.
PR	01-DEC-2000	2000US-0250160P.
PR	01-DEC-2000	2000US-0250391P.
PR	05-DEC-2000	2000US-0251030P.
PR	05-DEC-2000	2000US-0251988P.
PR	05-DEC-2000	2000US-0256719P.
PR	06-DEC-2000	2000US-0251479P.
PR	08-DEC-2000	2000US-0251856P.
PR	08-DEC-2000	2000US-0251868P.
PR	08-DEC-2000	2000US-0251869P.
PR	08-DEC-2000	2000US-0251989P.
PR	08-DEC-2000	2000US-0251990P.
PR	11-DEC-2000	2000US-0254097P.
PR	05-JAN-2001	2001US-0259678P.
PR	17-JAN-2001	2001US-0076486P.
PA		
XX	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PI	Rosen CA,	Ruben SM, Barash SC;
XX		
XX	WPI: 2003-743766/70.	
DR	N-PSDB;	ADE45654.
XX		
PT	New cardiovascular system related polynucleotides and polypeptides,	
PT	useful for preventing, treating, or ameliorating a medical condition,	
PT	such as cancer of cardiovascular tissues and cancer metastases.	
XX		
PS	Claim 11;	SEQ ID NO 1075; 262pp; English.
XX		
CC	The invention relates to human cardiovascular system related polypeptides	
CC	and the polynucleotides encoding them. The polypeptides, polynucleotides	
CC	and antibodies to the polypeptides are useful for diagnosing a	
CC	pathological condition or a susceptibility to a pathological condition,	
CC	for preventing, treating, or ameliorating a medical condition, such as	
CC	cancer of cardiovascular system tissues, proliferative disorders, foetal	
CC	and developmental abnormalities, haematopoietic disorders, diseases of	
CC	the immune system, AIDS, autoimmune diseases (e.g., Rheumatoid	
CC	arthritis), inflammation, allergies, neurological disorders (e.g.,	
CC	Alzheimer's disease, Parkinson's disease), cognitive disorders,	
CC	schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis,	
CC	diabetes, atherosclerosis, cardiovascular disorders, angiogenic	
CC	disorders, kidney disorders, gastrointestinal disorders, pregnancy-	
CC	related disorders, endocrine disorders and infections. The nucleic acids	
CC	are also useful for chromosome identification, radiation hybrid mapping	
CC	or long-range restriction mapping. The polypeptides and polynucleotides	
CC	may also be used as food additives or preservatives to increase or	

CC decrease storage capabilities, fat content or other nutritional  
 CC components. This sequence represents a human cardiovascular system  
 CC related polypeptide of the invention.  
 XX  
 SQ Sequence 67 AA;

Query Match 79.5%; Score 35; DB 7; Length 67;  
 Best Local Similarity 77.8%; Pred. No. 15;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9  
 |||||:  
 Db 7 FLOLLEPV 15

RESULT 6  
 ABB9785  
 ID ABB9785 standard; protein; 329 AA.

XX ABB9785;

XX 24-MAR-2003 (first entry)

XX Amino acid sequence of canine PXR ligand binding domain.

XX Pregnane X nuclear receptor; PXR; P450 3A4 monooxygenase;

KW drug metabolism.

XX Canine sp.

XX WO200294865-A1.

XX 28-NOV-2002.

XX 24-MAY-2002; 2002WO-US016445.

XX 24-MAY-2001; 2001US-0293380P.

XX (SMIK ) SMITHKLINE BEECHAM CORP.

XX Kliever SA, Maglich JM, Moore JT, Moore LB, Willson TM;

XX WPI; 2003-148458/14.

XX N-PSDB; AB23279.

XX New pregnane X nuclear receptor polypeptide that modulate P450 3A4 levels  
 PT or activity, useful in comparative pharmacology and for selecting  
 PT appropriate pre-clinical animal models predictive of human PXR activity.

XX Claim 1; Page 47-48; 56pp; English.

XX The present sequence is a pregnane X nuclear receptor (PXR) ligand  
 CC binding domain. PXR polypeptides modulate P450 3A4 monooxygenase levels  
 CC or activity. P450 3A4 catalyzes the metabolism of more than 60% of all  
 CC drugs that are in use, including steroids, immunosuppressive agents,  
 CC imidazole antimycotics, and macrolide antibiotics. The polypeptide is  
 CC useful in comparative pharmacology and for selecting appropriate pre-  
 CC clinical animal models predictive of human PXR activity

XX Sequence 329 AA;

Query Match 79.5%; Score 35; DB 6; Length 329;  
 Best Local Similarity 77.8%; Pred. No. 78;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9  
 |||||:  
 Db 210 FLOLLEPV 218

RESULT 7  
 ABP53033  
 ID ABP53033 standard; protein; 307 AA.

XX ABP53033;  
 AC  
 XX  
 DT 06-NOV-2002 (first entry)  
 XX  
 XX

C. elegans RNase H homologous protein sequence SEQ ID NO:3.

XX RNase H; antisense technology; inhibition.

XX Caenorhabditis elegans.

XX WO200264841-A1.

XX 22-AUG-2002.

XX 12-FEB-2002; 2002WO-US004243.

XX 12-FEB-2001; 2001US-00781712.

XX (ISIS-) ISIS PHARM INC.

XX Crooke ST, Lima WF, Wu H;

XX WPI; 2002-657606/70.

XX Use of a mammalian, particularly human, RNase H, for treating an animal  
 PT with a disease or condition associated with a human RNase H, for  
 PT inhibiting the expression of a protein, or for reducing cellular RNA via  
 PT antisense technology.

XX Example 1; Fig 1; 70pp; English.

XX The present invention describes a method for promoting the inhibition of  
 CC the expression of a protein comprising employing a mammalian RNase H  
 CC polypeptide so that cleavage of an RNA strand of an oligonucleotide-RNA  
 CC complex duplex occurs. Also described is a compound 8 to 50 nucleobases  
 CC in length targeted to the nucleic acid encoding the human RNase H1  
 CC polypeptide, where the compound specifically hybridizes with and inhibits  
 CC the expression of a human RNase H1 polypeptide. The compound, which is  
 CC an antisense oligonucleotide, is useful for inhibiting the expression of  
 CC a human RNase H1 polypeptide in cells or tissues, as well as for  
 CC treating an animal with a disease or condition associated with a human  
 CC RNase H1 polypeptide. The method is useful for inhibiting the expression  
 CC of a protein, particularly for reducing cellular RNA via antisense  
 CC technology. The present sequence represents a protein sequence given in  
 CC comparison with the human RNase H1 protein sequence, given in the  
 CC exemplification of the present invention

XX Sequence 307 AA;

Query Match 77.3%; Score 34; DB 5; Length 307;  
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9  
 |||||:  
 Db 227 FLOLSEPV 235

RESULT 8  
 AAY85935

ID AAY85935 standard; protein; 318 AA.

AC AAY85935;

XX 10-APR-2000 (first entry)

XX S. pneumoniae derived protein #144.

XX Treatment; prevention; disease; diagnosis; gene therapy; screening;  
 KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.

XX Streptococcus pneumoniae.

XX WO9806734-A1.  
XX 19-FEB-1998.  
XX 15-AUG-1997; 97WO-US014436.  
XX 16-AUG-1996; 96US-0024022P.  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
XX Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;  
PI Stodola RK;  
XX WPI; 1998-159452/14.  
DR N-PSDB; AA296294.  
XX Streptococcus pneumoniae proteins and related DNA - useful for screening  
PT compounds for antibacterial activity.  
PS Claim 5; Page 430-431; 640pp; English.  
XX This invention describes novel isolated Streptococcus pneumoniae  
CC polynucleotides (see AA296173-296494) and their encoded proteins (see  
CC AA296792-296182). The DNA, vectors and host cells described in the method  
CC of the invention are useful for the recombinant expression of the  
CC polypeptides. The polypeptides are useful for treatment or prevention of  
CC disease, or diagnosis of disease related to expression or activity of  
CC such a polypeptide. They can also be used to screen for compounds which  
CC interact with and inhibit or activate such a polypeptide. The  
CC polypeptides (or DNA encoding them, via gene therapy) are also useful for  
CC inducing an immunological response in a mammal. The antagonists are  
CC useful to inhibit such bacterial polypeptides. The polypeptides are  
CC particularly useful to identify antimicrobial compounds and antibiotics.  
CC They are also useful to determine their role in pathogenesis of  
CC infection, dysfunction and disease  
XX  
SQ Sequence 318 AA;  
Query Match 77.3%; Score 34; DB 2; Length 318;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLQILMEP 8  
Db 152 FLQVLM DP 159  
RESULT 9  
ABG61495  
ID ABG61495 standard; protein; 318 AA.  
XX  
XX ABG61495;  
XX 07-AUG-2003 (revised)  
DT 27-AUG-2002 (first entry)  
XX  
XX Iron uptake ABC transporter polypeptide #2.  
DE  
XX  
XX Iron uptake ABC transporter; Streptococcal iron transporter; Sit;  
KW antimicrobial; antibacterial; veterinary treatment; passive immunisation.  
XX  
XX Salmonella sp.  
OS  
XX  
XX WO200234773-A2.  
PN  
XX  
XX 02-MAY-2002.  
PD  
XX  
XX 26-OCT-2001; 2001WO-GB004749.  
PF  
XX  
XX 26-OCT-2000; 2000GB-00026231.  
PR 21-NOV-2000; 2000GB-00028345.  
PR 02-FEB-2001; 2001GB-00002666.  
PR

PR 02-MAY-2001; 2001US-0288118P.  
XX  
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.  
XX  
XX Holden DW, Brown JS;  
PI  
XX WPI; 2002-463352/49.  
DR  
XX N-PSDB; ABK83158.  
DR  
XX  
XX Novel Streptococcus pneumoniae iron uptake ABC transporter peptide,  
PT useful in screening assay for identifying antimicrobial drug and in  
PT diagnostic assay for detecting streptococcal microorganism.  
XX  
XX Disclosure; Page 51-52; 159pp; English.  
PS  
XX  
XX The invention relates to a Streptococcus pneumoniae iron uptake ABC  
CC transporter peptide called Streptococcal iron transporter (Sit) and the  
CC polynucleotide encoding it. The sequences are useful for therapeutic or  
CC diagnostic purposes, in screening assays for the identification of a  
CC antimicrobial drug, and in diagnostic assays for the detection of a  
CC streptococcal microorganism. The sequences are used for the manufacture  
CC of a medicament for the treatment or prevention of a condition associated  
CC with infection by S. pneumoniae or other gram positive bacteria,  
CC preferably for veterinary treatment, and in the production of monoclonal  
CC and polyclonal antibodies for use in passive immunisation. This sequence  
CC represents an S. pneumoniae iron uptake ABC transporter. (Updated on 07-  
CC AUG-2003 to correct OS field.)  
XX  
SQ Sequence 318 AA;  
Query Match 77.3%; Score 34; DB 5; Length 318;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLQILMEP 8  
Db 152 FLQVLM DP 159  
RESULT 10  
ABU02362  
ID ABU02362 standard; protein; 318 AA.  
XX  
XX ABU02362;  
XX  
XX 23-OCT-2003 (revised)  
DT 11-FEB-2003 (first entry)  
XX  
XX S. pneumoniae type 4 strain protein from coding region #1940.  
DE  
XX  
XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;  
KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;  
KW gene therapy; vaccine.  
XX  
XX Streptococcus pneumoniae; type 4 strain.  
OS  
XX  
XX WO200277021-A2.  
PN  
XX  
XX 03-OCT-2002.  
PD  
XX  
XX 27-MAR-2002; 2002WO-IB002163.  
PF  
XX  
XX 27-MAR-2001; 2001GB-00007658.  
PR  
XX  
XX (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
XX  
XX Masignani V, Tettelin H, Fraser C;  
PI  
XX  
XX WPI; 2003-040579/03.  
DR N-PSDB; ABX07652.  
XX  
XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,  
PT

useful as medicaments for treating or preventing a disease or infection due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or ear infection.

Claim 1; SEQ ID NO 3880; 56pp; English.

The invention relates to a protein comprising or having at least 50% identity to any of the 2469 amino acid sequences, identified in the specification (available on a computer readable format), or its fragment, expressed from 2469 of 2489 identified DNA coding regions from the Streptococcus pneumoniae type 4 strain genomic sequence appearing as AB56454. Also included are an antibody which binds one of the proteins, treating a patient by administering the protein, DNA or antibody (in a composition), a kit comprising first and second primers, which are the nucleic acid cited above or fragments between nucleotides 8-100 of a sequence not defined in the specification, for amplifying a target sequence contained within a Streptococcus nucleic acid sequence, where the first primer is substantially complementary to the target sequence and the second primer is substantially complementary to the complement of the target sequence, and where the parts of the primers having substantial complementarity define the termini of the target sequence to be amplified, assay comprising contacting a test compound with the protein, and determining whether the test compound binds to the protein and a Streptococcus pneumoniae bacterium, where one or more genes encoding the proteins has been rendered inactive. The proteins, nucleic acid molecules, antibody and compositions are useful as medicaments for treating or preventing a disease or infection due to streptococcus bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis media or ear infection. They are also useful in developing vaccines, diagnostics and antibiotics. The methods are useful for identifying immunodominant proteins. The present sequence is one of the 2469 proteins expressed by the identified coding regions from the genomic sequence. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at fcp.wipo.int/pub/published\_pct\_sequences. (Updated on 23-OCT-2003 to standardise OS field)

Sequence 318 AA;

Query Match 77.3%; Score 34; DB 6; Length 318;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMERP 8  
DB 152 FLOVIMDP 159

RESULT 11  
ADK46461  
ADK46461 standard; protein; 318 AA.

AC ADK46461;  
XX  
DT 20-MAY-2004 (first entry)

DE Streptococcus pneumoniae protein, Seq ID No 2976.

XX  
KM Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.

OS Streptococcus pneumoniae.

XX  
PN US6699703-B1.

XX  
PD 02-MAR-2004.

PF 26-MAY-2000; 2000US-00583110.

XX  
XX 02-JUL-1997; 97US-0051553P.  
PR 12-MAY-1998; 98US-0085131P.  
PR 30-JUN-1998; 98US-00107433.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.

XX  
PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Housewart CB;  
XX  
XX WPI; 2004-212399/20.  
DR N-PSDB; ADK43800.

XX  
PT New nucleic acid molecules and polypeptides useful for diagnosing,  
PT preventing and treating pathological conditions resulting from bacterial  
PT infection, e.g. Streptococcus pneumoniae infection, and in drug  
PT screening.

XX  
PS Disclosure, SEQ ID NO 2976; 301pp; English.

XX  
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids  
CC and polypeptides. The nucleic acids and proteins are useful for  
CC diagnosing, preventing and treating pathological conditions resulting  
CC from bacterial infection, such as S. pneumoniae infection. These may also  
CC be used for drug screening procedures. The present sequence represents a  
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
CC data for this patent did not appear in the printed specification but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html.

XX  
SQ Sequence 318 AA;

Query Match 77.3%; Score 34; DB 8; Length 318;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMERP 8  
DB 152 FLOVIMDP 159

RESULT 12  
AAV57148  
AAV57148 standard; protein; 2785 AA.

XX  
XX AAV57148;  
AC  
DT 28-FEB-2000 (first entry)

DE Human down-regulated in metastasis (DRIM) amino acid sequence.

XX  
KM Down-regulated in metastasis; DRIM; human; antimetastatic activity;  
KM antibody; tumour; treatment; therapy.

XX  
OS Homo sapiens.

OS  
PN W09960116-A1.  
XX  
PD 25-NOV-1999.

XX  
PF 17-MAY-1999; 99WO-EP003396.

XX  
PR 18-MAY-1998; 98EP-00303895.

XX  
PA (HOF) ROCHE DIAGNOSTICS GMBH.  
PA (ISIS-) ISIS INNOVATION LTD.

XX  
PI Weidle U, Tarin D;  
XX  
XX WPI; 2000-053296/04.  
DR N-PSDB; AAZ45136.

XX  
PT New polypeptide with antimetastatic activity, useful for therapeutic  
PT compositions for tumor therapy.

XX  
PS Claim 2; Page 42-49; 54pp; English.

XX  
XX This is the amino acid sequence of the human down-regulated in metastasis  
CC (DRIM) protein. The protein has antimetastatic activity. The DRIM protein  
CC is active in both its glycosylated and unglycosylated form, and can be



CC produced by recombinant technology in prokaryotic cells. DRIM mRNA is  
 CC strongly expressed in heart, skeletal muscle, pancreas, testis and ovary  
 CC tissues. The nucleotide and protein sequences can be used to create anti-  
 CC DRIM antibodies. The nucleic acids are useful in therapeutic  
 CC compositions, especially for treating tumours. They are also useful for  
 CC activating polynucleotides from the 5' untranslated region in gene  
 CC therapy

XX SQ Sequence 2785 AA;

Query Match 77.3%; Score 34; DB 3; Length 2785;  
 Best Local Similarity 77.8%; Pred. NO. 1.1e+03;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 FLOLMERP 9

DB 1056 FLDLLEPVP 1064

RESULT 13

ID ABP34618 standard; protein; 117 AA.

XX AC ABP34618;

XX DT 08-JUN-2002 (first entry)

XX DE Human ORF3591 protein, SEQ ID NO: 7182.

XX Human, ORF, open reading frame; ORFX; drug screening; diagnosis;  
 XX disease monitoring; cytokine; cell proliferation; cell differentiation;  
 XX immune modulation; haematopoiesis regulation; tissue growth;  
 XX angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;  
 XX thrombolytic; tumour inhibition; bodily characteristics; fertility;  
 XX behaviour; cancer; proliferative disorder; neurological disorder;  
 XX cardiovascular disease; immune system disorder; organ transplantation;  
 XX tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
 XX hypothyroidism; cholesterol ester storage disease; infection; vulnery;  
 XX vasotrophic; antipsoriatic; antidiabetic; cytosolic; nootropic;  
 XX neuroprotective; antithrombotic; anticoagulant; thrombolytic;  
 XX cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;  
 XX dermatological; analgesic; virucide; antibacterial; fungicide.

XX OS Homo sapiens.

XX OS WO200190366-A2.

XX PD 29-NOV-2001.

XX PF 24-MAY-2001; 2001WO-US017076.

XX PR 24-MAY-2000; 2000US-0206690P.

XX PA (CURA-) CURAGEN CORP.

XX PI Leach MD, Shinkets RA;

XX DR WPI; 2002-106200/14.

XX DR N-PSDB; ABN78644.

XX Novel human polypeptides and polynucleotides useful for diagnosing,  
 XX preventing and treating cardiovascular disease, neurodegenerative,  
 XX hyperproliferative disorders and disorders related to organ  
 XX transplantation.

XX CS Claim 10; Page 2042; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins  
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-  
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses  
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively  
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to  
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX

CC polynucleotides, the recombinant production of ORFX proteins, antibodies  
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and  
 CC polypeptides, methods of screening for modulators of ORFX expression or  
 CC activity, and methods of screening individuals for a predisposition to an  
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide  
 CC range of biological activities, such as cytokine, cell proliferation,  
 CC cell differentiation, immune modulation, haematopoiesis regulation,  
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/  
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,  
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,  
 CC and antiinfective activity, and may also be involved in the determination  
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,  
 CC nucleic acids and antibodies may be used in the treatment of cancers,  
 CC other proliferative disorders such as psoriasis and benign tumours,  
 CC neurological disorders such as epilepsy and Alzheimer's disease,  
 CC cardiovascular diseases, immune system disorders, disorders related to  
 CC organ transplantation, disorders of tissue growth and regeneration,  
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester  
 CC storage disease, and infectious diseases caused by viral, bacterial,  
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a  
 CC source of primers and probes, in the detection of ORFX genomic sequences  
 CC or transcripts, in the identification and cloning of homologous  
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX  
 CC nucleic acids may additionally be used to produce transgenic animals  
 CC which may be useful for studying the function and/or activity of ORFX  
 CC protein, and in drug screening. The ORFX proteins may also be used as  
 CC immunogens to generate specific antibodies, which are useful in the  
 CC diagnosis, treatment and monitoring of ORFX-associated diseases

XX SQ Sequence 117 AA;

Query Match 75.0%; Score 33; DB 5; Length 117;  
 Best Local Similarity 87.5%; Pred. NO. 70;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLOLMERP 8

DB 67 FLOLMERP 74

RESULT 14

ID AAG4343 standard; protein; 236 AA.

XX AC AAG4343;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 55533.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 XX hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.

XX OS Arabidopsis thaliana.

XX OS EP1033405-A2.

XX PN 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137727P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139482P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143628P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144684P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145376P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147995P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149375P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.

PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160982P.  
PR 25-OCT-1999; 99US-0161040P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 75.0%; Score 33; DB 3; Length 296;  
Best Local Similarity 87.5%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQILMEP 8  
DB 217 FLQILMEP 224

RESULT 15

ID AAG82085 standard; protein; 305 AA.

AC AAG82085;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:1264.

KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;  
endocarditis.

OS Staphylococcus epidermidis.

PN WO200134809-A2.

PD 17-MAY-2001.

PE 09-NOV-2000; 2000WO-US030782.

PR 09-NOV-1999; 99US-0164258P.

PA (GLAXO ) GLAXO GROUP LTD.

PI Kimmery WJ;

PS WPI; 2001-316495/33.

DR N-PSDB; AAH52935.

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
useful for vaccinating against infections, e.g. endocarditis.

ES Claim 18; Page 361; 2188pp; English.

CC AAH52304 to AAH5970 represent nucleic acids (I) encoding polypeptides  
(II), given in AAG81454 to AAG8120, from Staphylococcus epidermidis. (I)  
and (II) can have antibacterial activity and therefore can be used in  
vaccination. The nucleic acids (I) may be used to produce the S.  
epidermidis polypeptides (II) via the production of vectors containing

CC them which are used to produce hosts cells which express the  
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
CC used to vaccinate subjects and to raise antibodies against the bacteria.  
CC The polypeptides may also be used to assay for other inhibitors of their  
CC activity and therefore identify compounds that may be used for the  
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH5971 to  
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA  
CC polynucleotide sequences from the present invention. AAH5091 to AAH5098  
CC represent oligonucleotide sequences and primers which are used in the  
CC exemplification of the present invention. N.B. The present invention  
CC specifically claims all the polynucleotide sequences given in the  
CC sequence listing of the present specification, however the sequence  
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given  
CC in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present  
CC for SEQ ID NO:4455 to 4464

SO Sequence 305 AA;

Query Match 75.0%; Score 33; DB 4; Length 305;  
Best Local Similarity 75.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQILMEP 8  
DB 138 FLQILMEP 145

RESULT 16

ID ABM72286 standard; protein; 318 AA.

AC ABM72286;

DT 20-NOV-2003 (first entry)

DE Staphylococcus aureus protein #1526.

KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;  
enzymatic assay; antibiotic target.

OS Staphylococcus aureus.

PN WO200294868-A2.

PD 28-NOV-2002.

PE 27-MAR-2002; 2002WO-1B002637.

PR 27-MAR-2001; 2001GB-00007661.

PA (CHIR-) CHIRON SPA.

PI Maignani V, Mora M, Scarselli M;

PS WPI; 2003-120786/11.

DR N-PSDB; ACF73846.

PT New Staphylococcus aureus protein, useful as a vaccine for treating or  
preventing Staphylococcal infection, specifically an infection caused by  
S. aureus, e.g. sepsis.

ES Claim 1; SEQ ID NO 3052; 49pp; English.

CC The invention relates to novel genes and encoded proteins from  
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a  
CC nucleic acid encoding the protein, or an antibody to the protein, is  
CC useful as a pharmaceutical, particularly as a vaccine for treating or  
CC preventing infection due to Staphylococcus bacteria, specifically an  
CC infection caused by S. aureus. The composition is particularly useful for  
CC treating or preventing sepsis in a patient. The composition can also be  
CC used for diagnostics. The protein is also used in an assay for enzymatic  
CC studies and as a target for antibiotics. This sequence represents one of  
CC the novel S. aureus proteins of the invention

XX SQ Sequence 318 AA; 75.0%; Score 33; DB 6; Length 318;  
 Query Match Best Local Similarity 62.5%; Pred. No. 2e+02; Mismatches 5; Conservative 3; Indels 0; Gaps 0;  
 Oy 1 FLOLIMP 8  
 Db 151 FLOLIMP 158

RESULT 17  
 ABP40680  
 ID ABP40680 standard; protein; 324 AA.  
 AC ABP40680;  
 DT 24-JUL-2002 (first entry)  
 DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5525.  
 XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
 KW antibacterial; gene therapy.  
 OS Staphylococcus epidermidis.  
 PN US6380370-B1.  
 PD 30-APR-2002.  
 PF 13-AUG-1998; 98US-00134001.  
 PR 14-AUG-1997; 97US-0055779P.  
 PR 08-NOV-1997; 97US-0064964P.  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 PI Doucette-Stamm LA, Bush D;  
 DR WPI; 2002-381255/41.  
 DR N-PSDB; ABN933225.  
 PT Novel isolated nucleic acid encoding a Staphylococcus epidermis  
 PT polypeptide, useful for diagnosing and treating bacterial infections.  
 PS Disclosure; SEQ ID NO 5525; 267pp; English.  
 XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
 CC given in ABP95124 to ABP37960. The S. epidermidis sequences have  
 CC antibacterial activity and can be used in gene therapy. The sequences can  
 CC also be used in the diagnosis and treatment of bacterial infections,  
 CC particularly S. epidermidis infections. The sequences can be used to  
 CC screen for compounds able to interfere with the S. epidermidis life cycle  
 CC or inhibit S. epidermidis infection. N.B. The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from the USPTO web site  
 XX SQ Sequence 324 AA;  
 Query Match 75.0%; Score 33; DB 5; Length 324;  
 Best Local Similarity 75.0%; Pred. No. 2e+02; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 FLOLIMP 8  
 Db 157 FLOLIMP 164

RESULT 18  
 AAG44342  
 ID AAG44342 standard; protein; 345 AA.

XX AAG44342;  
 AC 18-OCT-2000 (first entry)  
 DT Arabidopsis thaliana protein fragment SEQ ID NO: 55532.  
 DE Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 OS Arabidopsis thaliana.  
 PN EP1033405-A2.  
 PD 06-SEP-2000.  
 PF 25-FEB-2000; 2000EP-00301439.  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 23-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-MAY-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0137222P.  
 PR 01-JUN-1999; 99US-0137528P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137528P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 14-JUN-1999; 99US-0138847P.  
 PR 16-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 17-JUN-1999; 99US-0139453P.  
 PR 18-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.



```

KW Rad23; maize; ATCC No: PTA-531; recombinant expression cassette; wheat;
XX transgenic plant; soybean; sunflower; sorghum; canola; modulator.
OS Zea mays.
XX
XX WOZ0001268-A1.
PN PD
XX 02-JUN-2000.
XX PF
XX 12-OCT-1999; 99WO-US042129.
XX PR
XX 23-NOV-1998; 98US-0109728P.
XX PA
XX (PION-) PIONEER HI-BRED INT INC.
XX MAHAJAN PB, Tagliani L;
XX DR
XX WPI: 2000-400078/34.
XX N-P8DB; AAD01231.
XX PT
XX Isolated nucleic acid encoding maize RAD23 protein is used to modulate
PT the levels of polypeptides in plant or in assays for identifying
PT compounds that bind to and/or increase/decrease enzymatic activity of
PT catalytically active polypeptides.
XX
XX Claim 11b; Page 78-79; 82pp; English.
PS
XX The present sequence is the maize Rad23 protein #2. It is isolated from a
CC Zea mays callus line, B73 callus tissue regenerated five days after
CC transfer of the callus from medium containing auxin to a medium devoid of
CC exogenous auxin. The cDNA is deposited under the ATCC No: PTA-531. Maize
CC Rad23 DNA sequence operably linked to a promoter, can be used to
CC construct a recombinant expression cassette. This expression cassette can
CC be used to generate a dicot or monocot transgenic plant e.g., maize,
CC soybean, sunflower, sorghum, canola, wheat, etc.. It can also be used to
CC modulate the levels of Rad23 polypeptide expression in a plant or in
CC assays to identify compounds, that bind to and/or modulate the enzymatic
CC activity of catalytically active polypeptides
XX
SQ Sequence 368 AA;

Query Match          75.0%; Score 33; DB 3; Length 368;
Best Local Similarity 87.5%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY      1 FROLMEP 8
        |||||
Db       291 FLOLMEP 298

RESULT 20
AAG44341
ID   AAG44341 standard; protein; 368 AA.
XX
XX   AAG44341;
AC
XX
DT    18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 55531.
XX
XX   Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX   Arabidopsis thaliana.
OS
XX
XX   EP1033405-A2.
XX
XX   06-SEP-2000.
PD
XX
PF    25-FEB-2000; 2000EP-00301439.
XX
PR    25-FEB-1999; 99US-0121825P.
XX

```

```

PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149910P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.

```

```

PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0158293P.
PR 13-OCT-1999; 99US-0158294P.
PR 13-OCT-1999; 99US-0158295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

```

```

Query Match 75.0%; Score 33; DB 3; Length 368;
Beet Local Similarity 87.5%; Pred. NO. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 FLOLMEP 8
Db 289 FLOLMEP 296

```

```

RESULT 21
AAW75700
ID AAW75700 standard; peptide; 409 AA.
XX
AC AAW75700;
XX
XX 29-OCT-1998 (first entry)
DT
XX
DE Vpr protein binding HHR23B amino acid sequence.
XX
XX Lentiviral infection; Vpr protein; HIV infection; cell stasis;
XX cell death; HHR23A; Rad23 protein.
OS
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 184..232
FT /note="highly conserved internal repeat domain"
FT Domain 360..385
FT /note="highly conserved internal repeat domain"
FT Domain 386..409
FT /note="highly conserved internal repeat domain"
XX
XX
XX W09835234-A1.
XX
XX 13-AUG-1998.
XX
XX 11-FEB-1998; 98WO-US003008.
XX
XX 11-FEB-1997; 97US-007979707.

```

XX (REGC ) UNIV CALIFORNIA.  
 PA Chen ISY, Jowett JBM, Withers-Ward E;  
 PI WPI; 1998-447375/38.  
 DR Identification of compounds binding the HIV-1 Vpr protein - that block  
 XX Vpr-mediated cell stasis, useful for treating HIV-1 or other lentiviral  
 PT infections.  
 PT  
 PS Disclosure; Fig 2; 63pp; English.  
 XX This represents the amino acid sequence of the HHR23B protein (a human  
 CC homologue of the S. cerevisiae Rad23 protein). This protein can bind to  
 CC the Vpr protein encoded by the HIV genome. The invention provides a  
 CC method of identifying an agent for use in treating lentiviral infections.  
 CC The method comprises contacting a cellular target of the Vpr protein with  
 CC the agent to be tested, and assessing the ability of the agent to block  
 CC interaction of the Vpr protein with the cellular target, where an agent  
 CC which blocks this interaction is an anti-lentiviral agent. Alternatively,  
 CC the agent contacts a cell expressing the Vpr protein under conditions  
 CC where the Vpr protein induces cell stasis in the absence of the agent.  
 CC Identification of the agent is then observed by blockade of Vpr-induced  
 CC cell stasis. The method allows the identification of compounds that block  
 CC Vpr-mediated cell stasis and ultimately cell death. The compounds can  
 CC thus be used in the treatment of HIV and other lentiviral infections  
 XX  
 SQ Sequence 409 AA;  
 XX  
 Query Match 75.0%; Score 33; DB 2; Length 409;  
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 FLOQLMEPV 9  
 Db 324 FLOQLMEPV 332  
 XX  
 RESULT 22  
 ID AAW68186 standard; protein; 409 AA.  
 AC AAW68186;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 29-OCT-1998 (first entry)  
 XX  
 DB Vpr binding protein HHR23B amino acid sequence.  
 XX  
 KM LentiVirus; Vpr protein; HIV infection; cell stasis; cell death; cancer;  
 KM HHR23B; Rad23 protein.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 184..232  
 FT Domain /note= "highly conserved internal repeat domain"  
 FT Domain 360..385  
 FT Domain /note= "highly conserved internal repeat domain"  
 FT Domain 386..409  
 FT Domain /note= "highly conserved internal repeat domain"  
 XX  
 PN WO9835032-A2.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PF 11-FEB-1998; 98WO-US003390.  
 XX  
 PR 11-FEB-1997; 97US-00798597.  
 PR 24-OCT-1997; 97US-00959279.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.

XX Chen ISY, Jowett JBM, Withers-Ward E, Stewart SA, Poon B;  
 PI Peigson J, Dieckmann T;  
 XX WPI; 1998-447229/38.  
 DR Arresting cell growth using lentivirus Vpr virion protein - used for  
 XX treatment of cancer and screening for agents that reduce Vpr binding,  
 PT e.g. anti-HIV agents.  
 PT  
 PS Disclosure; Fig 1; 71pp; English.  
 XX This represents the amino acid sequence of the HHR23B protein (a human  
 CC homologue of the S. cerevisiae Rad23 protein). This protein can bind to  
 CC the Vpr protein encoded by the HIV genome. This is used as a cellular  
 CC target in the method of the invention of identifying antitumour  
 CC therapeutic candidates. The invention provides a method for arresting the  
 CC growth of a cell by treatment with a Vpr lentivirus protein or its  
 CC analogue. Agents that reduce binding of Vpr to a cellular target are  
 CC useful for restoring growth. The antitumour agent identified is useful  
 CC for treating any type of cancer, since it induces cell stasis (blocks  
 CC development at the G2 stage) and death. The agents can also be used for  
 CC treating autoimmune diseases. (Updated on 25-MAR-2003 to correct PI  
 CC field.)  
 XX  
 SQ Sequence 409 AA;  
 XX  
 Query Match 75.0%; Score 33; DB 2; Length 409;  
 Best Local Similarity 66.7%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 FLOQLMEPV 9  
 Db 324 FLOQLMEPV 332  
 XX  
 RESULT 23  
 ID ABU07460 standard; protein; 409 AA.  
 AC ABU07460;  
 XX  
 DT 28-JAN-2003 (first entry)  
 XX  
 DE Protein differentially regulated in prostate cancer #63.  
 XX  
 KM Prostate cancer; gene expression; differential regulation;  
 KM molecular marker; drug target; cancer detection; cancer diagnosis;  
 KM cancer staging; cancer grading; cancer assessing; cancer monitoring.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200281638-A2.  
 XX  
 PD 17-OCT-2002.  
 XX  
 PF 08-APR-2002; 2002WO-US010824.  
 XX  
 PR 06-APR-2001; 2001US-0281731P.  
 PR 06-APR-2001; 2001US-0281732P.  
 XX  
 PA (ORIG-) ORIGENE TECHNOLOGIES INC.  
 XX  
 PI Sun Z, Jay G;  
 XX  
 DR WPI; 2003-058520/05.  
 DR N-PSDB; ABX10362.  
 XX  
 PT Novel genes which are differentially regulated in prostate cancer, useful  
 PT for diagnosing prostate cancer in prostate tissue sample and assessing  
 PT therapeutic or preventive intervention in prostate cancer patients.  
 XX





expression profile of a gene group comprising these genes. The expression levels or expression profiles produced by these genes are used as an indicator when screening for ischaemic condition-improving drugs or therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR primers for a mouse ischaemic condition related sequence, which are used in the exemplification of the present invention

CC Sequence 416 AA;

Query Match 75.0%; Score 33; DB 5; Length 416;  
Best Local Similarity 66.7%; Pred. No. 2.6e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEV 9  
Db 324 FLOLMPEV 332

RESULT 26

ABB66967  
ID ABB66967 standard; protein; 2764 AA.

XX ABB66967;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 27693.

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SBP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL11070.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

PS Disclosure; SEQ ID NO 27693; 21pp + Sequence listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

CC Sequence 2764 AA;

Query Match 75.0%; Score 33; DB 4; Length 2764;  
Best Local Similarity 55.6%; Pred. No. 1.8e+03;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMPEV 9

Db 1340 YLOMLPEPL 1348

RESULT 27

AAE05485  
ID AAE05485 standard; protein; 2802 AA.

XX AAE05485;

DT 24-SEP-2001 (first entry)

DE Drosophila neurofibromatosis type 1 (NF1) protein.

KW Neurofibromatosis type 1; NF1; cyclic AMP; therapy; benign tumour; malignant tumour; short stature; hyperpigmentation; white matter lesion; learning disability.

OS Drosophila sp.

XX Key Location/Qualifiers

FT Domain 1219..1580

FT /note="Catalytic GAP-related domain (GRD)"

XX US6261761-B1.

XX 17-JUL-2001.

PF 04-APR-2000; 2000US-00542331.

PR 24-MAR-1997; 97US-0041469P.

PR 24-MAR-1998; 98US-00046745.

PA (COLD-) COLD SPRING HARBOR LAB.

XX Zhong Y, Guo H;

DR WPI; 2001-450484/48.

PT Identifying compounds that induce production of cyclic AMP, by incubating compound with separate cellular preparations that differ by neurofibromatosis type I protein content and determining amount of cAMP produced.

PS Example 1; Col 11-26; 23pp; English.

CC The present invention relates to a method for identifying a compound that stimulate cyclic AMP formation comprising incubating compound with separate cellular preparations that differ by neurofibromatosis type I (NF1) protein content and determining amount of cyclic AMP produced. The compounds identified are used for the prevention of diseases associated with a defect in the NF1 protein. The diseases include neurofibromatosis type 1, benign tumours, malignant tumours, short stature, hyperpigmentation, white matter lesions in the brain and learning disabilities. The present sequence is Drosophila sp. neurofibromatosis type 1 (NF1) protein

CC Sequence 2802 AA;

Query Match 75.0%; Score 33; DB 4; Length 2802;  
Best Local Similarity 55.6%; Pred. No. 1.9e+03;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMPEV 9  
Db 1340 YLOMLPEPL 1348

RESULT 28

ABB63789  
ID ABB63789 standard; protein; 2802 AA.

XX ABB63789;

XX 26-MAR-2002 (first entry)  
 DT Drosophila melanogaster polypeptide SEQ ID NO 18159.  
 XX  
 DE Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical.  
 KW Drosophila melanogaster.  
 XX  
 OS WO200171042-A2.  
 XX  
 PN 27-SEP-2001.  
 XX  
 PD 23-MAR-2001; 2001WO-US009231.  
 XX  
 PF 23-MAR-2000; 2000US-0191637P.  
 XX  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PMD, Myers EW;  
 XX  
 DR MPI; 2001-656860/75.  
 XX  
 DR N-PSDB; ABL07892.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 18159; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 SQ Sequence 2802 AA;

Query Match 75.0%; Score 33; DB 4; Length 2802;  
 Best Local Similarity 55.6%; Pred. No. 1.9e+03;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQMLMEPV 9  
 DB 1340 YLQMLMEPL 1348

RESULT 29  
 ABB08077  
 ID ABB08077 standard; protein; 2802 AA.  
 XX  
 AC ABB08077;  
 XX  
 DT 10-SEP-2002 (first entry)  
 XX  
 DE Drosophila neurofibromatosis 1 (NFI) protein.  
 XX  
 KW NFI; pharmaceutical; memory; learning defect; neurofibromatosis 1;  
 KW neurotropic.  
 XX  
 OS Drosophila sp.  
 XX  
 PN US6365126-B1.  
 XX  
 PD 02-APR-2002.

PF 23-FEB-2000; 2000US-00510791.  
 XX  
 XX 24-MAR-1997; 97US-0041469P.  
 PR 24-MAR-1998; 98US-00046745.  
 XX  
 PA (COLD-) COLD SPRING HARBOR LAB.  
 XX  
 PI Zhong Y, Guo H, Tong J;  
 XX  
 DR MPI; 2002-487726/52.  
 XX  
 PT Screening an agent for ability to treat learning and short term memory  
 PT defects associated with defect in Neurofibromatosis 1 protein, by  
 PT administering agent to animal and assessing memory formation or learning  
 PT ability.  
 XX  
 PS Example 1; Fig 2A-D; 40pp; English.  
 XX  
 CC The invention relates to screening a pharmaceutical agent (PA) for its  
 CC ability to treat short term memory and learning defects in an animal,  
 CC associated with a defect in Neurofibromatosis 1 (NFI) protein. The method  
 CC involves administering PA to the animal, training the animal, assessing  
 CC memory formation or learning ability in the trained animal and comparing  
 CC it with that produced by training protocol in a control animal not  
 CC treated with PA. The identified agent is useful for treating learning or  
 CC short term memory defect associated with NFI protein in an animal. The  
 CC present sequence represents the Drosophila NFI protein sequence  
 CC  
 SQ Sequence 2802 AA;

Query Match 75.0%; Score 33; DB 5; Length 2802;  
 Best Local Similarity 55.6%; Pred. No. 1.9e+03;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQMLMEPV 9  
 DB 1340 YLQMLMEPL 1348

RESULT 30  
 ADA34524  
 ID ADA34524 standard; protein; 107 AA.  
 XX  
 AC ADA34524;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Acinetobacter baumannii protein #1685.  
 XX  
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;  
 KW plant biocontrol agent.  
 XX  
 OS Acinetobacter baumannii.  
 XX  
 PN US6562958-B1.  
 XX  
 PD 13-MAY-2003.  
 XX  
 PF 04-JUN-1999; 99US-00328352.  
 XX  
 PR 09-JUN-1998; 98US-0088701P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Breton G, Bush D;  
 XX  
 DR MPI; 2003-576092/54.  
 DR N-PSDB; ADA30398.  
 XX  
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
 PT for diagnosing a bacterial disease, as components of antibacterial  
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
 PT plants.

XX Example; SEQ ID NO 5811; 328pp; English.

PS The invention relates to isolated Acinetobacter baumannii nucleic acids.

CC The A. baumannii nucleic acids and polypeptides are useful as reagents

CC for diagnosing a bacterial disease, as components of antibacterial

CC vaccines, as targets for antibacterial drugs, to detect the presence of

CC A. baumannii and other Acinetobacter species in a sample, in screening

CC compounds for the ability to interfere with the A. baumannii life cycle

CC or to inhibit A. baumannii infection, and as biocontrol agents for A.

CC plants. The present sequence represents the amino acid sequence of an A.

CC baumannii protein.

XX

SEQ Sequence 107 AA;

Query Match 72.7%; Score 32; DB 6; Length 107;

Best Local Similarity 66.7%; Pred. No. 1e+02; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOQLMEPV 9

DB 2 YLEKLEMPV.10

RESULT 31

ADH86453

ID ADH86453 standard; protein; 195 AA.

XX

AC ADH86453;

XX

DT 22-APR-2004 (first entry)

XX

DE Enterococcus faecalis polypeptide #933.

XX

KM Enterococcus faecalis infection; transcription regulatory element;

XX

KM antibacterial.

XX

OS Enterococcus faecalis.

XX

PN US617156-B1.

XX

PD 09-SEP-2003.

XX

PF 13-AUG-1998; 98US-00134000.

XX

PR 15-AUG-1997; 97US-0055778P.

XX

PA (DOUC/) DOUCETTE-STAMM L A.

XX

PA (BUSH/) BUSH D.

XX

PI Doucette-Stamm LA, Bush D;

XX

DR WPI; 2003-895394/82.

XX

DR N-PSDB; ADH83048.

XX

PT New nucleic acid comprising a sequence encoding an Enterococcus faecalis

PT polypeptide, useful for preparing a composition for diagnosing or

PT treating E. faecalis infection.

XX

PS Disclosure; SEQ ID NO 4338; 193pp; English.

XX

CC The invention relates to Enterococcus faecalis polynucleotides and

CC polypeptides. The invention also relates to a recombinant expression

CC vector comprising a polynucleotide operably linked to a transcription

CC regulatory element, a cell comprising a recombinant vector, a method for

CC producing an E. faecalis polypeptide, an isolated nucleic acid comprising

CC a sequence not given in the specification, a recombinant vector

CC comprising the nucleic acid and a cell comprising the recombinant vector.

CC The polynucleotides can be used to detect the presence of E. faecalis in

CC a sample. The sequences are useful for preparing a composition for

CC diagnosing or treating Enterococcus faecalis infection. This sequence

CC represents an E. faecalis polypeptide of the invention.

XX

SEQ Sequence 195 AA;

Query Match 72.7%; Score 32; DB 7; Length 195;

Best Local Similarity 62.5%; Pred. No. 1.9e+02; Mismatches 3; Indels 0; Gaps 0;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOQLMEPV 8

DB 27 FLOQLMEPV 34

RESULT 32

AAV42689

ID AAV42689 standard; protein; 316 AA.

XX

AC AAV42689;

XX

DT 17-JAN-2000 (first entry)

XX

DE H1s6-pregnane X receptor (PXR) sequence.

XX

KM Human; nuclear receptor; pregnane X receptor; PXR; CYP3A4;

XX

KM cytochrome P-450 mono-oxygenase; drug interaction; hPXR.

XX

OS Synthetic.

XX

OS Homo sapiens.

XX

PN WO9948915-A1.

XX

PD 30-SEP-1999.

XX

PF 26-MAR-1999; 99WO-US006737.

XX

PR 27-MAR-1998; 98US-0079593P.

XX

PA (GLAX ) GLAXO GROUP LTD.

XX

PI Klierer SA, Willson TM;

XX

DR WPI; 1999-601202/51.

XX

PT New human pregnane X receptor, used to identify specific modulators and

PT agents that induce expression of cytochrome P-450 mono-oxygenase.

XX

PS Example 6; Page 37; 69pp; English.

XX

CC The invention provides an isolated human nuclear receptor (designated

CC pregnane X receptor, PXR) that binds to a cytochrome P-450 mono-oxygenase

CC (CYP) promoter. The hPXR is used to identify its specific modulators,

CC and compounds that induce CYP3A4 expression (i.e. to identify drug

CC interactions, since CYP3A4 is involved in many biotransformations of

CC drugs). The modulators are potentially useful for: associating particular

CC diseases and conditions with PXR and for treating such conditions.

CC Antibodies raised against hPXR can be used for determination and

CC purification of hPXR. The present sequence represents a histidine-6

CC tagged partial PXR (H1s6-PXR) sequence

XX

SEQ Sequence 316 AA;

Query Match 72.7%; Score 32; DB 2; Length 316;

Best Local Similarity 66.7%; Pred. No. 3.2e+02; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOQLMEPV 9

DB 197 FLOQLMEPV 205

RESULT 33

AAM50626

ID AAM50626 standard; protein; 316 AA.

XX

AC AAM50626;

XX 04-APR-2002 (first entry)  
 XX His6-tagged human pregnane X receptor.  
 DE  
 XX  
 XX  
 KM Pregnane X, receptor; hpxr; human; liver; CYP3A4;  
 KM cytochrome P450 monooxygenase; cirrhosis; cholangitis; hepatitis;  
 KM cholestasis; hepatotropic; antiinflammatory; therapy.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 OS Chimeric.  
 XX  
 XX Key Location/Qualifiers  
 FH Peptide 1..11  
 FT /label= His6  
 FT 12..316  
 FT Protein /label= hpxr  
 FT  
 PN WO200197856-A2.  
 XX  
 XX 27-DEC-2001.  
 PD  
 XX 21-JUN-2001; 2001WO-1B001629.  
 XX  
 XX 21-JUN-2000; 2000US-00598267.  
 XX  
 PA (GLAX ) GLAXO GROUP LTD.  
 XX  
 XX Kilewer SA, Jones SA, Willson TM,  
 DR WPI; 2002-139767/18.  
 XX  
 XX Compound that induces cytochrome P-450 monooxygenase 3A4 gene expression  
 PT for treating cholestatic liver disease comprising administering compound  
 PT identified by determining binding of test compound to human pregnane X  
 PT receptor.  
 XX  
 PS Example 6; Page 40; 63pp; English.  
 XX  
 XX The present sequence is that of a recombinant protein comprising an N-  
 CC terminal His6 tag followed by amino acids 130-434 of the novel human  
 CC pregnane X receptor (PXR, AAM50626). A DNA sequence encoding His6-PXR was  
 CC constructed in vector pRSER1. Plasmid His6-PXR/pRSER1 was cotransformed  
 CC with a plasmid encoding retinoic acid receptor-alpha (RXR-alpha) into  
 CC Escherichia coli BL21(DE3), and an His6-PXR/RXR-alpha complex was  
 CC obtained. hpxr is a novel member of the nuclear receptor superfamily. It  
 CC binds to a DNA response element in the cytochrome P450 monooxygenase  
 CC CYP3A4 gene promoter as a heterodimer with RXR, and is activated by  
 CC compounds known to modulate CYP3A4 expression. The invention provides  
 CC nucleic acids encoding hpxr, expression vectors, host cells, methods of  
 CC making the receptor, and methods of using the receptor or receptor-  
 CC encoding sequences to screen for compounds capable of modulating CYP  
 CC (e.g. CYP3A4) gene expression. Such compounds are useful for treating  
 CC cholestatic liver disease (claimed), such as primary biliary cirrhosis,  
 CC primary sclerosing cholangitis, autoimmune hepatitis with cholestatic  
 CC features, autoimmune cholangitis, cholestasis of pregnancy, paediatric  
 CC cholestatic syndromes, and drug-induced cholestasis  
 CC  
 XX  
 SQ Sequence 316 AA;  
 Query Match 72.7%; Score 32; DB 5; Length 316;  
 Best Local Similarity 66.7%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLOLIMBPV 9  
 DB 197 FQOLLBPM 205  
 RESULT 34  
 ADC95833  
 ID ADC95833 standard; protein; 323 AA.

XX AC. ADC95833;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE E. faecium protein sequence SEQ ID 5460.  
 XX  
 KM Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;  
 KM abdominal-pelvic infection.  
 XX  
 OS Enterococcus faecium.  
 OS  
 XX US6583275-B1.  
 XX  
 XX 24-JUN-2003.  
 PD  
 XX 30-JUN-1998; 98US-00107532.  
 XX  
 XX 02-JUL-1997; 97US-0051571P.  
 PR  
 XX 14-MAY-1998; 98US-0085598P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 XX Doucette-Stamm IA, Bush D;  
 PI  
 XX WPI; 2003-799836/75.  
 DR  
 XX N-PSDB; ADC92179.  
 XX  
 XX New isolated nucleic acid derived from Enterococcus faecium encoding an  
 PT Enterococcus faecium polypeptide useful for detection, prevention and  
 PT treatment of a pathological condition resulting from a bacterial  
 PT infection.  
 XX  
 PS Example 1; SEQ ID NO 5460; 243pp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid derived from  
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having  
 CC one of 10 fully defined sequences given in the (or comprising 40  
 CC sequential nucleotides chosen from any of the nucleic acids, its  
 CC complement or sequences hybridising to it). Also included are a  
 CC recombinant vector comprising the nucleic acid operably linked to  
 CC transcription regulatory element, a cell comprising the vector and a  
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are  
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.  
 CC The nucleic acid is useful for diagnosing pathological conditions  
 CC resulting from E. faecium bacterial infection (e.g. urinary tract  
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic  
 CC infection) and for screening drugs such as agonists and antagonists. The  
 CC nucleic acid is useful for recombinant production of Candida albicans -  
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions  
 CC and vaccines containing the nucleic acid are useful for preventing or  
 CC treating Enterococcus faecium infections. The present sequence represents  
 CC one if the disclosed E. faecium proteins.  
 CC  
 XX  
 SQ Sequence 323 AA;  
 Query Match 72.7%; Score 32; DB 7; Length 323;  
 Best Local Similarity 62.5%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLOLIMBP 8  
 DB 155 FLOVIMBP 162  
 RESULT 35  
 AAU78295  
 ID AAU78295 standard; protein; 329 AA.  
 XX  
 XX AAU78295;  
 AC  
 XX 18-JUN-2002 (first entry)  
 DT  
 XX

DE Human Pregnane X Receptor (PXR) ligand binding domain (LDB) protein.  
 XX Human; pregnane X; receptor; PXR; ligand binding domain; LDB; xenobiotic.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 XX WO200218420-A2.  
 XX PD 07-MAR-2002.  
 XX PF 17-AUG-2001; 2001WO-EP009488.  
 XX PR 28-AUG-2000; 2000EP-00118634.  
 XX PA (LION-) LION BIOSCIENCE AG.  
 XX PI Albers M, Ellwanger S, Koegl M, Loeser E;  
 XX DR WPI; 2002-292195/33.  
 XX DR N-PSDB; ABK47576.  
 XX PT New nucleic acids and cofactors of the pregnane X nuclear receptor (PXR),  
 XX PT which the nucleic acid encode, useful for screening agonists or  
 XX PT antagonists of PXR, and for determining a subject's response to  
 XX PT xenobiotic substances or drugs.  
 XX PS Claim 13; Fig 6; 102pp; English.  
 XX CC The present invention relates to a new nucleic acid molecule and its  
 XX CC encoded polypeptide. The nucleic acid codes for a cofactor of the  
 XX CC pregnane X nuclear receptor (PXR). The polypeptide encoded by the nucleic  
 XX CC acid comprises 225 amino acids or 293 amino acids fully defined in the  
 XX CC specification. The nucleic acid is useful for making vectors and  
 XX CC transforming cells, both of which are ultimately useful for producing the  
 XX CC cofactor proteins. The nucleic acids may also be used for determining a  
 XX CC subject's response to xenobiotic substances or drugs. The proteins or  
 XX CC complexes are useful for screening substances that bind the proteins or  
 XX CC complexes, particularly agonists or antagonists of PXR. The present amino  
 XX CC acid sequence represents the human PXR ligand binding domain (LBD) of the  
 XX CC invention  
 XX SQ Sequence 329 AA;  
 XX  
 XX Query Match 72.7%; Score 32; DB 5; Length 329;  
 XX Best Local Similarity 66.7%; Pred. No. 3.3e+02;  
 XX Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLOLMEPV 9  
 Db 210 FQQLLEPM 218  
 XX  
 XX RESULT 36  
 XX ID AAY42691  
 XX AAY42691 standard; protein; 414 AA.  
 XX AC AAY42691;  
 XX DT 17-JAN-2000 (first entry)  
 XX DE Human pregnane X receptor (hPXR).  
 XX KW Human; nuclear receptor; pregnane X receptor; PXR; CYP; CYP3A4;  
 XX KW Cytochrome P-450 mono-oxygenase; drug interaction; hPXR.  
 XX OS Homo sapiens.  
 XX OS WO9948915-A1.  
 XX PD 30-SEP-1999.  
 XX PF 26-MAR-1999; 99WO-US006737.  
 XX

PR 27-MAR-1998; 98US-0079593P.  
 XX (GLAXO) GLAXO GROUP LTD.  
 XX PI Kiewer SA, Willson TM;  
 XX DR WPI; 1999-601202/51.  
 XX DR N-PSDB; AAZ07997.  
 XX PT New human pregnane X receptor, used to identify specific modulators and  
 XX PT agents that induce expression of cytochrome P-450 mono-oxygenase.  
 XX PS Claim 4; Fig 1A-D; 69pp; English.  
 XX CC The invention provides an isolated human nuclear receptor (designated  
 XX CC pregnane X receptor, PXR) that binds to a cytochrome P-450 mono-oxygenase  
 XX CC (CYP) promoter. The hPXR is used to identify its specific modulators,  
 XX CC and compounds that induce CYP3A4 expression (i.e. to identify drug  
 XX CC interactions, since CYP3A4 is involved in many biotransformations of  
 XX CC drugs). The modulators are potentially useful for: associating particular  
 XX CC diseases and conditions with PXR and for treating such conditions.  
 XX CC Antibodies raised against hPXR can be used for determination and  
 XX CC purification of hPXR. The present sequence represents the hPXR  
 XX SQ Sequence 414 AA;  
 XX  
 XX Query Match 72.7%; Score 32; DB 2; Length 414;  
 XX Best Local Similarity 66.7%; Pred. No. 4.2e+02;  
 XX Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLOLMEPV 9  
 Db 315 FQQLLEPM 323  
 XX  
 XX RESULT 37  
 XX ID AAY15931  
 XX AAY15931 standard; protein; 434 AA.  
 XX AC AAY15931;  
 XX DT 04-AUG-1999 (first entry)  
 XX DE A human intranuclear receptor protein.  
 XX KW Human; intranuclear receptor protein; drug development; diagnosis;  
 XX KW treatment.  
 XX OS Homo sapiens.  
 XX OS JP1127872-A.  
 XX PD 18-MAY-1999.  
 XX PF 07-AUG-1998; 98JP-00224172.  
 XX PR 11-AUG-1997; 97JP-00230335.  
 XX PA (NISB) JAPAN TOBACCO INC.  
 XX DR WPI; 1999-350330/30.  
 XX DR N-PSDB; AAX59966.  
 XX PT New intranuclear receptor protein - useful for drug development and  
 XX PT diagnosis and treatment of disease.  
 XX PS Claim 1; Page 15-16; 38pp; Japanese.  
 XX CC The present sequence represents a human intranuclear receptor protein.  
 XX CC The nucleic acid sequence was isolated from a human adult cDNA library  
 XX CC using a swallfish ANO23 derived probe. The protein can be used for the  
 XX CC development of drugs and diagnosis and treatment of various diseases  
 XX

SQ Sequence 434 AA;  
 Query Match 72.7%; Score 32; DB 2; Length 434;  
 Best Local Similarity 66.7%; Pred. No. 4.4e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLOLMPEPV 9  
 |||:|:  
 Db 315 FQQLLEPM 323  
 RESULT 38  
 AAY21799  
 ID AAY21799 standard; protein; 434 AA.  
 XX  
 AC AAY21799;  
 XX  
 DT 14-SEP-1999 (first entry)  
 XX  
 DE Human steroid and xenobiotic receptor (SXR).  
 XX  
 KW Nuclear receptor; SXR; steroid and xenobiotic receptor; RXR; human;  
 KW retinoid X receptor; P450 gene; steroid hormone; steroid metabolism;  
 KW phytoestrogen; calcium-channel blocker; steroid toxicity; tuberculosis;  
 KW breast cancer; osteoporosis; Cushing syndrome; virilism; hirsutism;  
 KW polycystic ovarian disease; cancer; colorectal; prostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 227  
 FT /label= unknown  
 FT /note= "encoded by ACN"  
 XX  
 FM MO9935246-A1.  
 XX  
 PD 15-JUL-1999.  
 XX  
 PF 08-JAN-1999; 99WO-US000490.  
 XX  
 PR 09-JAN-1998; 98US-00005266.  
 XX  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX  
 PI Evans RM, Blumberg B;  
 XX  
 DR MPI; 1999-419349/35.  
 DR N-PSDB; AAX89080.  
 XX  
 PT New steroid and xenobiotic receptor, used to identify modulators for  
 PT controlling metabolism of steroid and xenobiotics, e.g. reducing their  
 PT toxicity.  
 XX  
 PS Claim 4; Fig 1A; 83pp; English.  
 CC  
 CC The invention relates to a novel nuclear receptor polypeptide, designated  
 CC SXR (steroid and xenobiotic receptor). SXR (i) forms a heterodimer with  
 CC retinoid X receptor (RXR), (ii) binds to a direct or inverted repeat  
 CC response element motif based on the half-site AGTTCA, (iii) activates  
 CC transcription through response elements present in steroid-inducible P450  
 CC genes, in response to a wide variety of natural and synthetic steroid  
 CC hormones and (iv) is prominently expressed in liver and intestine. SXR  
 CC regulates expression of catabolic enzymes, in response to many different  
 CC steroids, and thus affects metabolism. SXR is a broad specificity, low-  
 CC affinity receptor for reducing excessive levels of steroids in the  
 CC circulation. (Antagonists of SXR are used to regulate metabolism of  
 CC steroids particularly phytoestrogens or calcium-channel blockers, to  
 CC reduce steroid toxicity in subjects being treated with steroids, e.g. in  
 CC cases of tuberculosis (treated with rifampin and related compounds),  
 CC breast cancer (treated with tamoxifen, raloxifen etc.) or osteoporosis  
 CC (treated with Vitamin K), or to slow metabolism of therapeutic steroids.  
 CC Also, modulating endogenous SXR is used to treat disease, particularly an  
 CC agonist is used where endogenous steroid levels are excessive (e.g.

CC Cushing syndrome; virilism and hirsutism in women; polycystic ovarian  
 CC disease; 11 beta-, 17- or 21-hydroxylase deficiency; 3 beta-hydroxysteroid  
 CC dehydrogenase deficiency, or breast, colorectal or prostatic cancer),  
 CC while antagonists are used where endogenous steroid levels are too low.  
 CC Cells that express SXR are used to identify compounds likely to be  
 CC involved in undesirable drug interactions. Antibodies specific for SXR  
 CC are used in immunohistochemical testing for studying distribution/  
 CC expression density of SXR, also for diagnosis and therapeutically as  
 CC antagonists. The present sequence represents SXR polypeptide  
 CC  
 SQ Sequence 434 AA;  
 Query Match 72.7%; Score 32; DB 2; Length 434;  
 Best Local Similarity 66.7%; Pred. No. 4.4e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLOLMPEPV 9  
 |||:|:  
 Db 315 FQQLLEPM 323  
 RESULT 39  
 AAB84417  
 ID AAB84417 standard; protein; 434 AA.  
 XX  
 AC AAB84417;  
 XX  
 DT 22-AUG-2001 (first entry)  
 XX  
 DE Amino acid sequence of a steroid-activated nuclear receptor.  
 XX  
 KW Steroid-activated nuclear receptor; steroid and xenobiotic receptor; SXR;  
 KW retinoid X receptor; RXR; transcription; response element;  
 KW steroid inducible P450 gene; steroid hormone; Cushing's syndrome;  
 KW virilism; hirsutism; polycystic ovarian syndrome; hypertension.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 227  
 FT /note= "unspecified residue encoded by ACN"  
 XX  
 FM WO200142290-A2.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PF 08-DEC-2000; 2000WO-US033473.  
 XX  
 PR 09-DEC-1999; 99US-00458366.  
 XX  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX  
 PI Evans RM, Blumberg B, Xie W;  
 XX  
 DR MPI; 2001-381637/40.  
 DR N-PSDB; AAH25489.  
 XX  
 PT Novel steroid-activated nuclear receptor useful as sensor for xenobiotic  
 PT compounds and/or steroids and whose modulators are useful for modulating  
 PT metabolism of steroid or xenobiotic compounds.  
 XX  
 PS Disclosure; Fig 1A; 64pp; English.  
 CC  
 CC The present sequence is a human steroid-activated nuclear receptor,  
 CC termed steroid and xenobiotic receptor (SXR). The SXR polypeptide is  
 CC capable of forming a heterodimer with retinoid X receptor (RXR),  
 CC activating transcription through response elements found in steroid  
 CC inducible P450 genes in response to a variety of natural and synthetic  
 CC steroid hormones and prominently expressed in liver and intestine. SXR  
 CC binds to a direct or inverted repeat response element motif based on the  
 CC half site AGTTCA. SXR is useful for identifying compounds which are  
 CC agonists or which activate the receptor. The compounds identified are  
 CC useful for treating a wide variety of conditions such as Cushing's

CC syndrome, virilism and hirsutism, androgen excess due to polycystic  
CC ovarian syndrome and enzymatic defects which leads to accumulation of  
CC steroids, resulting in hypertension and aberrant development of secondary  
CC sexual characteristics in both sexes. Transgenic animals which express  
CC human SXR serve as models for human response to various agents which  
CC potentially impact P450-dependent metabolic processes  
XX

SQ Sequence 434 AA;

Query Match 72.7%; Score 32; DB 4; Length 434;

Best Local Similarity 66.7%; Pred. No. 4.4e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9  
| | | | |  
Db 315 FQQLLEPM 323

RESULT 40

ID AAU78294 standard; protein; 434 AA.

XX AAU78294;

DT 18-JUN-2002 (first entry)

DS Human Pregnane X Receptor (PXR) protein.

KW Human; pregnane X; receptor; PXR; xenobiotic.

OS Homo sapiens.

Key Location/Qualifiers

FT Region 56..434  
/note="Encoded by ABK47575"

PN WO200218420-A2.

PD 07-MAR-2002.

PF 17-AUG-2001; 2001WO-EP009488.

PR 28-AUG-2000; 2000EP-00118634.

PA (LION-) LION BIOSCIENCE AG.

PI Albers M, Ellwanger S, Koegl M, Loeser E;

DR N-PSDB; ABK47575.

DR WPI; 2002-292195/33.

PT New nucleic acids and cofactors of the pregnane X nuclear receptor (PXR),  
PT which the nucleic acid encode, useful for screening agonists or  
PT antagonists of PXR, and for determining a subject's response to  
PT xenobiotic substances or drugs.

PS Claim 13; Fig 5; 102pp; English.

XX The present invention relates to a new nucleic acid molecule and its  
XX encoded polypeptide. The nucleic acid codes for a cofactor of the  
XX pregnane X nuclear receptor (PXR). The polypeptide encoded by the nucleic  
XX acid comprises 225 amino acids or 293 amino acids fully defined in the  
XX specification. The nucleic acid is useful for making vectors and  
XX transforming cells, both of which are ultimately useful for producing the  
XX cofactor proteins. The nucleic acids may also be used for determining a  
XX subject's response to xenobiotic substances or drugs. The proteins or  
XX complexes are useful for screening substances that bind the proteins or  
XX complexes, particularly agonists or antagonists of PXR. The present amino  
XX acid sequence represents the human PXR protein of the invention

SQ Sequence 434 AA;

Query Match 72.7%; Score 32; DB 5; Length 434;

Best Local Similarity 66.7%; Pred. No. 4.4e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9  
| | | | |  
Db 315 FQQLLEPM 323

Search completed: January 12, 2005, 20:06:46  
Job time : 209.6 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2005, 20:16:00 ; Search time 93.6 Seconds  
(without alignments)  
34.739 Million cell updates/sec

Title: US-09-870-216C-3

Perfect score: 44

Sequence: 1 FLQGLMEPV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCF\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCFUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	9	11	US-09-870-216C-3
2	44	100.0	9	13	US-10-017-327-3
3	35	79.5	67	9	US-09-764-869-1075
4	35	79.5	67	14	US-10-091-504-1075
5	35	79.5	67	15	US-10-227-577-1075
6	35	79.5	329	16	US-10-478-551-2
7	35	79.5	363	15	US-10-424-599-170467
8	35	79.5	363	15	US-10-424-599-219691
9	34	77.3	79	17	US-10-425-115-149203
10	34	77.3	180	16	US-10-437-963-114660
11	34	77.3	274	16	US-10-437-963-132856
12	34	77.3	307	9	US-09-799-848-3
13	34	77.3	307	11	US-09-781-7128-3

14	34	77.3	307	17	US-10-679-761-8	Sequence 8, Appl1
15	34	77.3	518	16	US-10-415-478A-5	Sequence 5, Appl1
16	34	77.3	304	15	US-10-425-114-45949	Sequence 45949, A
17	34	77.3	770	16	US-10-437-963-163658	Sequence 163658, A
18	34	77.3	798	17	US-10-425-115-200455	Sequence 200455, A
19	33	75.0	80	17	US-10-425-115-319897	Sequence 319897, A
20	33	75.0	117	11	US-09-864-408A-7182	Sequence 7182, Ap
21	33	75.0	138	17	US-10-425-115-223947	Sequence 223947, A
22	33	75.0	174	17	US-10-425-115-223948	Sequence 223948, A
23	33	75.0	193	16	US-10-767-701-35283	Sequence 35283, A
24	33	75.0	349	17	US-10-425-115-223949	Sequence 223949, A
25	33	75.0	368	9	US-09-805-550-4	Sequence 4, Appl1
26	33	75.0	1224	16	US-10-437-963-138501	Sequence 138501, A
27	32	72.7	55	17	US-10-425-115-307557	Sequence 307557, A
28	32	72.7	73	16	US-10-767-701-56147	Sequence 56147, A
29	32	72.7	79	17	US-10-425-115-358291	Sequence 358291, A
30	32	72.7	104	16	US-10-767-701-62832	Sequence 62832, A
31	32	72.7	169	17	US-10-425-115-360733	Sequence 360733, A
32	32	72.7	189	15	US-10-424-599-196742	Sequence 196742, A
33	32	72.7	198	15	US-10-424-599-196742	Sequence 196742, A
34	32	72.7	269	15	US-10-424-599-234629	Sequence 234629, A
35	32	72.7	272	15	US-10-425-114-53064	Sequence 53064, A
36	32	72.7	276	17	US-10-425-115-367551	Sequence 367551, A
37	32	72.7	290	17	US-10-775-678-54	Sequence 54, Appl1
38	32	72.7	334	16	US-10-437-963-108157	Sequence 108157, A
39	32	72.7	362	17	US-10-739-930-9255	Sequence 9255, Ap
40	32	72.7	420	17	US-10-425-115-365824	Sequence 365824, A
41	32	72.7	434	10	US-09-143-828-2	Sequence 2, Appl1
42	32	72.7	434	10	US-09-227-718-2	Sequence 2, Appl1
43	32	72.7	434	10	US-09-840-008-2	Sequence 2, Appl1
44	32	72.7	434	14	US-10-081-555C-2	Sequence 2, Appl1
45	32	72.7	434	15	US-10-226-997-2	Sequence 2, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-870-216C-3  
; Sequence 3, Application US/09870216C  
; Publication No. US20040138135A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
; FILE REFERENCE: 68126881210100  
; CURRENT APPLICATION NUMBER: US/09/870,216C  
; CURRENT FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/209,391  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/226,256  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: 60/257,008  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-870-216C-3

Query Match 100.0%; Score 44; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0;

QY 1 FLQGLMEPV 9  
Db 1 FLQGLMEPV 9

RESULT 2  
US-10-017-327-3  
; Sequence 3, Application US/10017327

```
; Publication No. US20020155471A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Nicolette
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND
; FILE REFERENCE: G2 2101.20
; CURRENT APPLICATION NUMBER: US/10/017,327
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-327-3

Query Match      100.0%; Score 44; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLOLMEPV 9
      |||||
Db      1 FLOLMEPV 9

RESULT 3
; Sequence 1075, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1075
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (65)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-1075

Query Match      79.5%; Score 35; DB 9; Length 67;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLOLMEPV 9
      |||||
Db      7 FLOLMEPV 15

RESULT 4
; US-10-091-504-1075
; Sequence 1075, Application US/10091504
; Publication No. US2003005908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 1075
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-091-504-1075

Query Match      79.5%; Score 35; DB 14; Length 67;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 FLOLMEPV 9
      |||||
Db      7 FLOLMEPV 15

RESULT 5
; US-10-227-577-1075
; Sequence 1075, Application US/10227577
; Publication No. US2004000575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1075
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-227-577-1075

Query Match      79.5%; Score 35; DB 15; Length 67;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

OY 1 FLOLLMEPV 9  
||||: |||  
DB 7 FLOLLMEPV 15

RESULT 6  
US-10-478-551-2  
; Sequence 2, Application US/10478551  
; Publication No. US20040171811A1  
; GENERAL INFORMATION:  
; APPLICANT: SmithKline Beecham Corporation  
; APPLICANT: Steven Anthony Kijewer  
; APPLICANT: Jodi Marie Maglich  
; APPLICANT: John Tomlin Moore  
; APPLICANT: Linda Becker Moore  
; APPLICANT: Timothy Mark Willson  
; TITLE OF INVENTION: NONHUMAN PREGNANT X RECEPTOR SEQUENCES  
; TITLE OF INVENTION: FOR USE IN COMPARATIVE PHARMACOLOGY  
; FILE REFERENCE: PUS855W0  
; CURRENT APPLICATION NUMBER: US/10/478,551  
; CURRENT FILING DATE: 2003-11-21  
; PRIOR APPLICATION NUMBER: 60/293,380  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FASTSBQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 329  
; TYPE: PRT  
; ORGANISM: Canine  
US-10-478-551-2

Query Match 79.5%; Score 35; DB 16; Length 329;  
Best Local Similarity 77.8%; Pred. No. 1.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLOLLMEPV 9  
||||: |||  
DB 210 FLOLLMEPV 218

RESULT 7  
US-10-424-599-170467  
; Sequence 170467, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 170467  
; LENGTH: 363  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1) ..(363)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_124948C.1.pep  
US-10-424-599-170467

Query Match 79.5%; Score 35; DB 15; Length 363;  
Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLOLLMEPV 9  
||||: |||

DB 286 FLOLLMEPV 294

RESULT 8  
US-10-424-599-219691  
; Sequence 219691, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 219691  
; LENGTH: 363  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_40409C.1.pep  
US-10-424-599-219691

Query Match 79.5%; Score 35; DB 15; Length 363;  
Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLOLLMEPV 9  
||||: |||  
DB 286 FLOLLMEPV 294

RESULT 9  
US-10-425-115-349203  
; Sequence 349203, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 349203  
; LENGTH: 79  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_81639C.1.pep  
US-10-425-115-349203

Query Match 77.3%; Score 34; DB 17; Length 79;  
Best Local Similarity 66.7%; Pred. No. 42;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLLMEPV 9  
||||: |||  
DB 66 YLOLLMEPV 74

RESULT 10  
US-10-437-963-114660  
; Sequence 114660, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yinhua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Bardazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53221)B  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 114660  
LENGTH: 180  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)-(180)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_18329C.1.pep  
US-10-437-963-114660

Query Match 77.3%; Score 34; DB 16; Length 180;  
Best Local Similarity 77.8%; Pred. No. 99;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 169 FCOLMHPV 177

RESULT 11  
US-10-437-963-132856  
Sequence 132856, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yinhua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Bardazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53221)B  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 132856  
LENGTH: 274  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)-(274)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_34784C.1.pep  
US-10-437-963-132856

Query Match 77.3%; Score 34; DB 16; Length 274;  
Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 184 FCOLMHPV 192

RESULT 12  
US-09-799-848-3  
Sequence 3, Application US/09799848  
Patent No. US20010044145A1  
GENERAL INFORMATION:  
APPLICANT: Monia, Brett  
APPLICANT: Cook, Phillip  
APPLICANT: Crooke, Stanley  
APPLICANT: Wu, Hongjiang  
APPLICANT: Lima, Walter  
TITLE OF INVENTION: METHODS OF USING MAMMALIAN RNASE H AND COMPOSITIONS THEREOF  
FILE REFERENCE: ISPH-0521  
CURRENT FILING DATE: US/09/799,848  
CURRENT FILING DATE: 2001-03-05  
PRIOR APPLICATION NUMBER: US 09/343,809  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: US 09/684,254  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: US 09/203,716  
PRIOR FILING DATE: 1998-12-02  
PRIOR APPLICATION NUMBER: US 60/067,458  
PRIOR FILING DATE: 1997-12-04  
PRIOR APPLICATION NUMBER: US 09/453,514  
PRIOR FILING DATE: 1999-12-01  
PRIOR APPLICATION NUMBER: US 09/144,611  
PRIOR FILING DATE: 1998-08-31  
PRIOR APPLICATION NUMBER: US 08/861,306  
PRIOR FILING DATE: 1997-04-21  
PRIOR APPLICATION NUMBER: US 08/244,993  
PRIOR FILING DATE: 1994-06-21  
PRIOR APPLICATION NUMBER: US 07/814,961  
PRIOR FILING DATE: 1991-12-24  
PRIOR APPLICATION NUMBER: US 09/462,280  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: PCT/US98/13966  
PRIOR FILING DATE: 1998-07-06  
PRIOR APPLICATION NUMBER: US 08/889,296  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: US 08/411,734  
PRIOR FILING DATE: 1995-04-03  
PRIOR APPLICATION NUMBER: US 08/007,996  
PRIOR FILING DATE: 1993-10-21  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3  
LENGTH: 307  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-09-799-848-3

Query Match 77.3%; Score 34; DB 9; Length 307;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 227 FLOLSEPV 235

RESULT 13  
US-09-781-712B-3  
Sequence 3, Application US/09781712B  
Publication No. US20040180433A1  
GENERAL INFORMATION:  
APPLICANT: Crooke, Stanley T  
APPLICANT: Lima, Walter  
APPLICANT: Wu, Hongjiang  
TITLE OF INVENTION: Methods of Using Mammalian RNase H and Compositions Thereof  
FILE REFERENCE: ISPH-0520  
CURRENT FILING DATE: US/09/781,712B  
CURRENT FILING DATE: 2001-02-12  
PRIOR APPLICATION NUMBER: US 60/067,458

PRIOR FILING DATE: 1997-12-04  
 PRIOR APPLICATION NUMBER: US 09/203,716  
 PRIOR FILING DATE: 1998-12-02  
 PRIOR APPLICATION NUMBER: US 09/343,809  
 PRIOR FILING DATE: 1999-06-30  
 PRIOR APPLICATION NUMBER: US 09/684,254  
 PRIOR FILING DATE: 2000-10-06  
 NUMBER OF SEQ ID NOS: 39  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 3  
 LENGTH: 307  
 TYPE: PRT  
 ORGANISM: Caenorhabditis elegans  
 US-09-781-712B-3

Query Match 77.3% Score 34; DB 11; Length 307;  
 Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOQLMEPV 9  
 DB 227 FLOQLSVEPV 235

RESULT 14  
 US-10-679-761-8  
 Sequence 8, Application US/10679761  
 Publication No. US20040248145A1  
 GENERAL INFORMATION:  
 APPLICANT: Isis Pharmaceuticals, Inc.  
 APPLICANT: Crooke, Stanley T.  
 APPLICANT: Lima, Walter  
 APPLICANT: Wu, Hongjiang  
 TITLE OF INVENTION: Methods of Using Mammalian Rhase H and Compositions Thereof  
 FILE REFERENCE: ISPH-0790  
 CURRENT FILING DATE: US/10/679,761  
 PRIOR FILING DATE: 2003-10-06  
 PRIOR APPLICATION NUMBER: US 10/358,439  
 PRIOR FILING DATE: 2003-02-03  
 PRIOR APPLICATION NUMBER: US 09/992,738  
 PRIOR FILING DATE: 2001-11-14  
 PRIOR APPLICATION NUMBER: US 09/781,712  
 PRIOR FILING DATE: 2001-02-12  
 PRIOR APPLICATION NUMBER: US 09/861,205  
 PRIOR FILING DATE: 2001-05-18  
 PRIOR APPLICATION NUMBER: US 09/684,254  
 PRIOR FILING DATE: 2000-10-06  
 PRIOR APPLICATION NUMBER: US 09/343,809  
 PRIOR FILING DATE: 1999-06-30  
 PRIOR APPLICATION NUMBER: US 09/203,716  
 PRIOR FILING DATE: 1998-12-02  
 PRIOR APPLICATION NUMBER: US 60/067,458  
 PRIOR FILING DATE: 1997-12-04  
 PRIOR APPLICATION NUMBER: US 60/248,950  
 PRIOR FILING DATE: 2000-11-15  
 PRIOR APPLICATION NUMBER: US 60/497,412  
 PRIOR FILING DATE: 2003-08-21  
 NUMBER OF SEQ ID NOS: 104  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO 8  
 LENGTH: 307  
 TYPE: PRT  
 ORGANISM: Caenorhabditis elegans  
 US-10-679-761-8

Query Match 77.3% Score 34; DB 17; Length 307;  
 Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOQLMEPV 9  
 DB 227 FLOQLSVEPV 235

RESULT 15  
 US-10-415-478A-5  
 Sequence 5, Application US/10415478A  
 Publication No. US20040116661A1  
 GENERAL INFORMATION:  
 APPLICANT: Jeremy Stuart Brown  
 APPLICANT: David William Holden  
 TITLE OF INVENTION: Streptococcal Genes  
 FILE REFERENCE: GJE-6571  
 CURRENT APPLICATION NUMBER: US/10/415,478A  
 CURRENT FILING DATE: 2003-12-28  
 PRIOR APPLICATION NUMBER: PCT/GB01/04749  
 PRIOR FILING DATE: 2001-10-26  
 PRIOR APPLICATION NUMBER: 0026231.1  
 PRIOR FILING DATE: 2000-10-26  
 PRIOR APPLICATION NUMBER: 00028345.7  
 PRIOR FILING DATE: 2000-11-21  
 PRIOR APPLICATION NUMBER: 0102666.5  
 PRIOR FILING DATE: 2001-02-02  
 PRIOR APPLICATION NUMBER: 60/288,118  
 PRIOR FILING DATE: 2001-05-02  
 NUMBER OF SEQ ID NOS: 91  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 5  
 LENGTH: 318  
 TYPE: PRT  
 ORGANISM: Salmonella pneumoniae  
 US-10-415-478A-5

Query Match 77.3% Score 34; DB 16; Length 318;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOQLMEPV 8  
 DB 152 FLOQLMDEPV 159

RESULT 16  
 US-10-425-114-45949  
 Sequence 45949, Application US/10425114  
 Publication No. US20040034888A1  
 GENERAL INFORMATION:  
 APPLICANT: Liu, Jingdong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven E  
 APPLICANT: Tabaska, Jack E  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(5313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO 45949  
 LENGTH: 504  
 TYPE: PRT  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: LIB3594-024-A4\_FLI.pep  
 US-10-425-114-45949

Query Match 77.3% Score 34; DB 15; Length 504;  
 Best Local Similarity 77.8%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOQLMEPV 9  
 DB 67 FLOQLMDEPV 75

## RESULT 17

US-10-437-963-163658  
; Sequence 163658, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Bardazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 163658  
; LENGTH: 770  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(770)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_62633C.1.pep  
US-10-437-963-163658

Query Match 77.3%; Score 34; DB 16; Length 770;  
Best Local Similarity 77.8%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9  
| | | | |  
DB 187 FCOLMHPV 195

## RESULT 18

US-10-425-115-200455  
; Sequence 200455, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 200455  
; LENGTH: 798  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(798)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_114401C.1.pep  
US-10-425-115-200455

Query Match 77.3%; Score 34; DB 17; Length 798;  
Best Local Similarity 77.8%; Pred. No. 4.6e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9  
| | | | |

DB 224 FCOLMHPV 232

## RESULT 19

US-10-425-115-319897  
; Sequence 319897, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 319897  
; LENGTH: 80  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_54815C.1.pep  
US-10-425-115-319897

Query Match 75.0%; Score 33; DB 17; Length 80;  
Best Local Similarity 66.7%; Pred. No. 68;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9  
| | | | |  
DB 54 FLNLMHPI 62

## RESULT 20

US-09-864-408A-7182  
; Sequence 7182, Application US/09864408A  
; Publication No. US20040009474A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Shinkets, Richard A.  
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Enco  
; FILE REFERENCE: 21402-012  
; CURRENT APPLICATION NUMBER: US/09/864,408A  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/206,690  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 908  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7182  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-864-408A-7182

Query Match 75.0%; Score 33; DB 11; Length 117;  
Best Local Similarity 87.5%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEPV 8  
| | | | |  
DB 67 FLOLMEPV 74

## RESULT 21

US-10-425-115-223947  
; Sequence 223947, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 223947  
LENGTH: 138  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_135825C.1.pep  
US-10-425-115-223947

Query Match  
Best Local Similarity 75.0%; Score 33; DB 17; Length 138;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEP 8  
Db 61 FLOLMEP 68

RESULT 22  
US-10-425-115-223948  
Sequence 223948, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 223948  
LENGTH: 174  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(174)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_135826C.1.pep  
US-10-425-115-223948

Query Match  
Best Local Similarity 75.0%; Score 33; DB 17; Length 174;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEP 8  
Db 97 FLOLMEP 104

RESULT 23  
US-10-767-701-35283  
Sequence 35283, Application US/10767701  
Publication No. US20040172684A1  
GENERAL INFORMATION:  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/767,701  
CURRENT FILING DATE: 2004-01-29

NUMBER OF SEQ ID NOS: 63128  
SEQ ID NO 35283  
LENGTH: 193  
TYPE: PRT  
ORGANISM: Sorghum bicolor  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(193)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C6308\_1.pep  
US-10-767-701-35283

Query Match  
Best Local Similarity 75.0%; Score 33; DB 16; Length 193;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEP 8  
Db 116 FLOLMEP 123

RESULT 24  
US-10-425-115-223949  
Sequence 223949, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 223949  
LENGTH: 349  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(349)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_135827C.1.pep  
US-10-425-115-223949

Query Match  
Best Local Similarity 75.0%; Score 33; DB 17; Length 349;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEP 8  
Db 297 FLOLMEP 304

RESULT 25  
US-09-805-550-4  
Sequence 4, Application US/09805550  
Patent No. US20020026045A1  
GENERAL INFORMATION:  
APPLICANT: Mahajan, Pramod B.  
APPLICANT: Tasliani, Laura  
TITLE OF INVENTION: Rad23 Genes and Uses Thereof  
FILE REFERENCE: 0964D  
CURRENT APPLICATION NUMBER: US/09/805,550  
CURRENT FILING DATE: 2001-03-13  
PRIOR FILING DATE: 1999-10-06  
PRIOR APPLICATION NUMBER: 60/109,728  
PRIOR FILING DATE: 1998-11-23

NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 368  
TYPE: PRT  
ORGANISM: Zea mays  
US-09-805-550-4

Query Match 75.0%; Score 33; DB 9; Length 368;  
Best Local Similarity 87.5%; Pred. No. 3.3e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMMEPV 8  
DB 291 FLOLMMEPV 298

RESULT 26  
US-10-437-963-138501  
Sequence 138501, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 138501  
LENGTH: 1224  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_39884C.1.pep  
US-10-437-963-138501

Query Match 75.0%; Score 33; DB 16; Length 1224;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMMEPV 9  
DB 257 FLOLMMEPV 265

RESULT 27  
US-10-425-115-307557  
Sequence 307557, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 307557  
LENGTH: 55  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:

OTHER INFORMATION: Clone ID: MRT4577\_43561C.1.pep  
US-10-425-115-307557

Query Match 72.7%; Score 32; DB 17; Length 55;  
Best Local Similarity 66.7%; Pred. No. 73;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLMMEPV 9  
DB 37 FLOLMMEPV 45

RESULT 28  
US-10-767-701-56147  
Sequence 56147, Application US/10767701  
Publication No. US20040172684A1  
GENERAL INFORMATION:  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53535)B  
CURRENT APPLICATION NUMBER: US/10/767,701  
CURRENT FILING DATE: 2004-01-29  
NUMBER OF SEQ ID NOS: 63128  
SEQ ID NO 56147  
LENGTH: 73  
TYPE: PRT  
ORGANISM: Sorghum bicolor  
FEATURE:  
OTHER INFORMATION: Clone ID: 30937978.pep  
US-10-767-701-56147

Query Match 72.7%; Score 32; DB 16; Length 73;  
Best Local Similarity 62.5%; Pred. No. 98;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMMEPV 8  
DB 55 FLOLMMEPV 62

RESULT 29  
US-10-425-115-358291  
Sequence 358291, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 358291  
LENGTH: 79  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_8992C.1.pep  
US-10-425-115-358291

Query Match 72.7%; Score 32; DB 17; Length 79;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMMEPV 9  
DB 64 FLOLMMEPV 72



RESULT 30  
US-10-767-701-62832  
; Sequence 62832, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 62832  
; LENGTH: 104  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 18066026.pep  
US-10-767-701-62832

Query Match 72.7%; Score 32; DB 16; Length 104;  
Best Local Similarity 87.5%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLMPEV 9  
|||  
Db 64 LQLCPEV 71

RESULT 31  
US-10-425-115-360733  
; Sequence 360733, Application US/10425115  
; Publication No. US2004021472A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 360733  
; LENGTH: 169  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1) ..(169)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT457\_92170C.1.pep  
US-10-425-115-360733

Query Match 72.7%; Score 32; DB 17; Length 169;  
Best Local Similarity 75.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLMPEV 9  
|||  
Db 56 LQMLPEV 63

RESULT 32  
US-10-424-599-245981  
; Sequence 245981, Application US/10424599  
; Publication No. US20040031072A1

; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 245981  
; LENGTH: 189  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_64151C.1.pep  
US-10-424-599-245981

Query Match 72.7%; Score 32; DB 15; Length 189;  
Best Local Similarity 55.6%; Pred. No. 2.6e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQMLPEV 9  
|||  
Db 22 FMRLLQPV 30

RESULT 33  
US-10-424-599-196742  
; Sequence 196742, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 196742  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_19685C.1.pep  
US-10-424-599-196742

Query Match 72.7%; Score 32; DB 15; Length 198;  
Best Local Similarity 87.5%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLMPEV 9  
|||  
Db 132 LQMLPEV 139

RESULT 34  
US-10-424-599-234629  
; Sequence 234629, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599

```

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 234629
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(269)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53899C.1.pcp
US-10-424-599-234629

Query Match          72.7%; Score 32; DB 15; Length 269;
Best Local Similarity 55.6%; Pred. No. 3.8e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9
DB 59 FKRLLIQPV 67

RESULT 35
US-10-425-114-53064
; Sequence 53064, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53064
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3158-004-D4_FLI.pcp
US-10-425-114-53064

Query Match          72.7%; Score 32; DB 15; Length 272;
Best Local Similarity 66.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9
DB 113 FYQLVPEPL 121

RESULT 36
US-10-425-115-367551
; Sequence 367551, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 367551
```

```

; LENGTH: 276
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_98378C.1.pcp
US-10-425-115-367551

Query Match          72.7%; Score 32; DB 17; Length 276;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9
DB 117 FYQLVPEPL 125

RESULT 37
US-10-775-678-54
; Sequence 54, Application US/10775678
; Publication No. US20040229250A1
; GENERAL INFORMATION:
; APPLICANT: Transkaryotic Therapies, Inc.
; APPLICANT: von Figura, Kurt
; APPLICANT: Schmidt, Bernhard
; APPLICANT: Dierks, Thomas
; APPLICANT: Hartlein, Michael W.
; APPLICANT: Cosma, Maria P.
; APPLICANT: Ballabio, Andrea
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MULTIPLE SULFATASE DEFICIENCY AND
; FILE REFERENCE: 0403
; CURRENT APPLICATION NUMBER: US/10/775,678
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,747
; PRIOR FILING DATE: 2003-02-11
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Anopheles gambiae
US-10-775-678-54

Query Match          72.7%; Score 32; DB 17; Length 290;
Best Local Similarity 77.8%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLMEPV 9
DB 84 FOQLSSEPV 92

RESULT 38
US-10-437-963-108157
; Sequence 108157, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 108157
; LENGTH: 334
```

TYPE: PRT  
 ORGANISM: Oryza sativa  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT4530\_12438C.1.pcp  
 US-10-437-963-108157

Query Match 72.7%; Score 32; DB 16; Length 334;  
 Best Local Similarity 66.7%; Pred. No. 4.7e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
 |||:|:  
 Db 180 FYQLVEPL 188

RESULT 39  
 US-10-739-930-9255  
 ; Sequence 9255, Application US/10739930  
 ; Publication No. US20040216190A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kovalic, David K.  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
 ; FILE REFERENCE: 38-21(53377)B  
 ; CURRENT APPLICATION NUMBER: US/10/739,930  
 ; CURRENT FILING DATE: 2003-12-18  
 ; NUMBER OF SEQ ID NOS: 11088  
 ; SEQ ID NO 9255  
 ; LENGTH: 362  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: GLYMA-23APR03-C573\_22.p  
 US-10-739-930-9255

Query Match 72.7%; Score 32; DB 17; Length 362;  
 Best Local Similarity 87.5%; Pred. No. 5.1e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LQLEMEPV 9  
 |||:|:  
 Db 132 LQLEMEPV 139

RESULT 40  
 US-10-425-115-365824  
 ; Sequence 365824, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53322)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 365824  
 ; LENGTH: 420  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(420)  
 ; OTHER INFORMATION: unsure at all Xaa locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_96797C.1.pcp  
 US-10-425-115-365824

Query Match 72.7%; Score 32; DB 17; Length 420;  
 Best Local Similarity 66.7%; Pred. No. 6e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 FLOLMEPV 9  
 |||:|:  
 Db 262 FYQLVEPL 270

Search completed: January 12, 2005, 20:48:24  
 Job time : 102 secs

THIS PAGE BLANK (COPY)

THE

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:46:14 ; Search time 16.4 Seconds

(without alignments)  
52.802 Million cell updates/sec

Title: US-09-870-216C-3

Perfect score: 44

Sequence: 1 FLQULMEPV 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	35	79.5	379 2	RAD23 protein, iso
2	34	77.3	318 2	iron-compound ABC
3	34	77.3	318 2	hypothetical prote
4	34	77.3	1254 2	hypothetical prote
5	33	75.0	113 2	F309.1 protein - A
6	33	75.0	113 2	hypothetical prote
7	33	75.0	409 2	hypothetical prote
8	33	75.0	908 2	hypothetical prote
9	33	75.0	2764 2	gene col intron 1
10	33	75.0	2802 2	neurofibromin - fr
11	33	75.0	2802 2	neurofibromin - fr
12	33	72.7	882 2	neurofibromin - fr
13	32	72.7	480 2	neurofibromin - fr
14	32	72.7	480 2	neurofibromin - fr
15	32	72.7	504 2	neurofibromin - fr
16	32	72.7	536 2	neurofibromin - fr
17	32	72.7	643 2	neurofibromin - fr
18	32	72.7	687 2	neurofibromin - fr
19	32	72.7	1101 2	neurofibromin - fr
20	32	72.7	1510 2	neurofibromin - fr
21	31	70.5	142 1	gonadotropin II be
22	31	70.5	142 1	gonadotropin II be
23	31	70.5	255 2	conserved hypoch
24	31	70.5	342 2	conserved hypoch
25	31	70.5	401 2	conserved hypoch
26	31	70.5	456 2	conserved hypoch
27	31	70.5	539 2	conserved hypoch
28	31	70.5	639 2	conserved hypoch
29	31	70.5	730 2	conserved hypoch
30	31	70.5	730 2	conserved hypoch
31	31	70.5	730 2	conserved hypoch
32	31	70.5	730 2	conserved hypoch
33	31	70.5	730 2	conserved hypoch
34	31	70.5	730 2	conserved hypoch
35	31	70.5	730 2	conserved hypoch
36	31	70.5	730 2	conserved hypoch
37	31	70.5	730 2	conserved hypoch
38	31	70.5	730 2	conserved hypoch
39	31	70.5	730 2	conserved hypoch
40	31	70.5	730 2	conserved hypoch
41	31	70.5	730 2	conserved hypoch
42	31	70.5	730 2	conserved hypoch
43	31	70.5	730 2	conserved hypoch
44	31	70.5	730 2	conserved hypoch
45	31	70.5	730 2	conserved hypoch

30	31	70.5	921 2	probable helicase
31	31	70.5	921 2	probable helicase
32	31	70.5	921 2	probable helicase
33	31	70.5	921 2	probable helicase
34	31	70.5	921 2	probable helicase
35	31	70.5	921 2	probable helicase
36	31	70.5	921 2	probable helicase
37	31	70.5	921 2	probable helicase
38	31	70.5	921 2	probable helicase
39	31	70.5	921 2	probable helicase
40	31	70.5	921 2	probable helicase
41	31	70.5	921 2	probable helicase
42	31	70.5	921 2	probable helicase
43	31	70.5	921 2	probable helicase
44	31	70.5	921 2	probable helicase
45	31	70.5	921 2	probable helicase

#### ALIGNMENTS

##### RESULT 1

FLQULMEPV 9

C:Species: Dactylis glomerata (carrot)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 16-Aug-2004

C:Accession: T14337

R:Sturm, A.; Leinhardt, S.

Plant J. 13, 815-821, 1998

A:Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.

A:Reference number: Z17989; MUID:98345997; PMID:9681019

A:Accession: T14337

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-379 <STU>

A:Cross-references: UNIPROT:O03991; EMBL:Y12014; NID:G1914684; PIDD:CAA72742.1; PIDD:G191

A:Experimental source: subspecies Queen Anne's lace, isolate W001C

C:Genetics:

A:Gene: RAD23-2

C:Superfamily: ubiquitin homology

Query Match 79.5%; Score 35; DB 2; Length 379;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLQULMEPV 9

Db 302 FLQULMEPV 310

##### RESULT 2

E95218

iron-compound ABC transporter, permease protein SPI870 [imported] - Streptococcus pneumo

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004

C:Accession: E95218

R:Terrell, H.; Nelson, K.B.; Paulsen, I.T.; Eisele, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel, I

son, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: E95218

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-318 <KUR>

A:Cross-references: UNIPROT:Q97NY1; GB:AB005672; PIDD:AAK75942.1; PIDD:G1497373; GSPDB:GN

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SPI870

Query Match 77.3%; Score 34; DB 2; Length 318;

Best Local Similarity 75.0%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMEP 8  
|||:|:|:  
Db 152 FLOVLMDP 159

## RESULT 3

hypothetical protein fatc [imported] - Streptococcus pneumoniae (strain R6)

C/Species: Streptococcus pneumoniae

C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004

C/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.; ansen, N.F.; Hughes, B.; Huzar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marzilli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A/Reference number: A86141; PMID:21016719; PMID:11130712

A/Accession: H86296

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-113 <STO>

A/Cross-references: UNIPROT:Q9SA20; GB:AE005172; NID:G4966345; PIDN:AA034676.1; GSPDB:GNC

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;

ansen, N.F.; Hughes, B.; Huzar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marzilli, Rizzo,

M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A/Reference number: A86141; PMID:21016719; PMID:11130712

A/Accession: H86296

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-113 <STO>

A/Cross-references: UNIPROT:Q9SA20; GB:AE005172; NID:G4966345; PIDN:AA034676.1; GSPDB:GNC

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A/Cross-references: UNIPROT:Q8DNJ4; GB:AE007317; PIDN:AL00488.1; PID:G15459360; GSPDB:G

C/Genetics:

A/Map position: 1

A/Reference number: A97872; PMID:21429245; PMID:11544234

A/Accession: C98082

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-318 <KUR>

A:Cross-references: UNIPROT:P54727; EMBL:D21090; NID:g498147; PIDD:BA04652.1; PID:d1005  
C:Superfamily: ubiquitin homology  
F:1-80/Domain: ubiquitin homology <UBH>

Query Match Best Local Similarity 75.0%; Score 33; DB 2; Length 409;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
Db 324 FLOLMPEPV 332

#### RESULT 8

S07649  
Gene: cot1 intron 1 protein - Neurospora crassa mitochondrion  
C:Species: mitochondrion Neurospora crassa  
C>Date: 29-Jan-1993 #sequence\_revision 29-Jan-1993 #text\_change 07-Dec-1999  
C/Accession: S07649  
R/Field: D.J.; Sommerfield, A.; Saville, B.J.; Collins, R.A.  
Nucleic Acids Res. 17, 9087-9099, 1989  
A>Title: A group II intron in the Neurospora mitochondrial cot gene: nucleotide sequence  
A/Reference number: S07649; MUID:90067912; PMID:2531370  
A/Accession: S07649  
A/Molecule type: DNA  
A/Residues: 1-908 <FIE>  
A:Cross-references: EMBL:X14669  
A/Experimental source: strain Adipodome  
C/Genetics:  
A/Genome: mitochondrion  
A/Genetic code: SGC3  
C:Superfamily: yeast mitochondrion ox13 intron 1 protein  
C/Keywords: mitochondrion

Query Match Best Local Similarity 75.0%; Score 33; DB 2; Length 908;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLMPEPV 9  
Db 369 LQLMPEPV 376

#### RESULT 9

T13949  
neurofibromin - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C/Accession: T13949  
R/The, I.; Hannigan, G.E.; Cowley, G.S.; Reginald, S.; Zhong, Y.; Gusella, J.F.; Harihar  
Science 276, 791-794, 1997  
A>Title: Rescue of a Drosophila Nf1 mutant phenotype by protein kinase A.  
A/Reference number: Z17826; MUID:97277221; PMID:9115203  
A/Accession: T13949  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-2764 <THE>  
A:Cross-references: UNIPROT:O01399; EMBL:L26502; NID:g1929432; PID:g1929433; PIDD:AA589  
C/Genetics:  
A/Genes: NF1  
A:Cross-references: FlyBase:FBgn0015269

Query Match Best Local Similarity 75.0%; Score 33; DB 2; Length 2764;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
Db 1340 YLOMLPEPL 1348

#### RESULT 10

T13945

neurofibromin - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C/Accession: T13945  
R/Bernards, A.  
submitted to the EMBL Data Library, April 1997  
A/Reference number: Z17824  
A/Accession: T13945  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-2802 <BER>  
A:Cross-references: UNIPROT:O01397; EMBL:L26500; NID:g1929428; PID:g1929429; PIDD:AA589  
C/Genetics:  
A/Genes: NF1  
A:Cross-references: FlyBase:FBgn0015269

A/Introns: 19/3; 67/3; 196/1; 243/2; 441/2; 548/3; 733/1; 912/2; 1443/1; 1645/2; 2430/3;  
Query Match Best Local Similarity 75.0%; Score 33; DB 2; Length 2802;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
Db 1340 YLOMLPEPL 1348

#### RESULT 11

T13947  
neurofibromin - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C/Accession: T13947  
R/The, I.; Hannigan, G.E.; Cowley, G.S.; Reginald, S.; Zhong, Y.; Gusella, J.F.; Harihar  
Science 276, 791-794, 1997  
A>Title: Rescue of a Drosophila Nf1 mutant phenotype by protein kinase A.  
A/Reference number: Z17826; MUID:97277221; PMID:9115203  
A/Accession: T13947  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-2802 <THE>  
A:Cross-references: UNIPROT:O01398; EMBL:L26501; NID:g1929430; PID:g1929431; PIDD:AA589  
C/Genetics:  
A/Genes: NF1  
A:Cross-references: FlyBase:FBgn0015269

Query Match Best Local Similarity 75.0%; Score 33; DB 2; Length 2802;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
Db 1340 YLOMLPEPL 1348

#### RESULT 12

T14336  
RAD23 protein, isoform I - carrot  
C/Species: Daucus carota (carrot)  
C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 16-Aug-2004  
C/Accession: T14336  
R/Sturm, A.; Leinhardt, S.  
Plant J. 13, 815-821, 1998  
A>Title: Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in yeast.  
A/Reference number: Z17989; MUID:98345997; PMID:9661019  
A/Accession: T14336  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-382 <STU>  
A:Cross-references: UNIPROT:O03990; EMBL:Y12013; NID:g1914682; PIDD:CAA72741.1; PID:g191  
A/Experimental source: subspecies Queen Anne's lace, isolate W001C  
C/Genetics:  
A/Genes: RAD23-1  
C:Superfamily: ubiquitin homology

Query Match 72.7%; Score 32; DB 2; Length 382;  
 Best Local Similarity 66.7%; Pred. No. 49;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
 DB 311 FLOLMPEPV 319

## RESULT 13

AB2042  
 xanthine/uracil permease family protein VC2712 [imported] - *Vibrio cholerae* (strain N16)  
 C/Species: *Vibrio cholerae*  
 C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C/Accession: AB2042  
 R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Charidson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bais, S.; Qin, H.; Dragoti, I.; Sellers, F. I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A/Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 A/Reference number: AB2035; MUID:20406833; PMID:10952301  
 A/Accession: AB2042  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-480 <HEI>  
 A/Cross-references: UNIPROT:Q9KMM0; GB:AE004337; GB:AE003852; NID:9657307; PIDN:AAF9585  
 C/Genetics:  
 A/Map position: 1  
 C/Superfamily: hypothetical protein b2882

Query Match 72.7%; Score 32; DB 2; Length 480;  
 Best Local Similarity 66.7%; Pred. No. 62;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
 DB 371 FLOLMPEPV 379

## RESULT 14

T04076  
 protoporphyrinogen oxidase (EC 1.3.3.4) IX, mitochondrial [validated] - common tobacco  
 C/Species: *Nicotiana tabacum* (common tobacco)  
 C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C/Accession: T04076  
 R/Leremontova, I.; Kruse, E.; Mock, H.P.; Grimm, B.  
 Proc. Natl. Acad. Sci. U.S.A. 94, 8895-8900, 1997  
 A/Title: Cloning and characterization of a plastidial and a mitochondrial isoform of tobacco  
 A/Reference number: Z15186; MUID:97385200; PMID:9238074  
 A/Accession: T04076  
 A/Status: preliminary; translated from GB/EMBL/DD8J  
 A/Molecule type: mRNA  
 A/Residues: 1-504 <LER>  
 A/Cross-references: UNIPROT:O24164; EMBL:Y13466; NID:92370334; PIDN:CAA73866.1; PID:9237  
 A/Experimental source: strain SRI  
 C/Genetics:  
 A/Gene: ppxII  
 C/Function:  
 A/Description: EC 1.3.3.4 [validated, MUID:97385200]  
 A/Pathway: tetrapyrrole synthesis  
 C/Keywords: mitochondrion; oxidoreductase

Query Match 72.7%; Score 32; DB 2; Length 504;  
 Best Local Similarity 62.5%; Pred. No. 65;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLOLMPEPV 9  
 DB 128 FLOLMPEPV 135

RESULT 15  
 F86251  
 hypothetical protein [imported] - *Arabidopsis thaliana*  
 C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C/Accession: F86251

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marzall, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A/Reference number: AB6141; MUID:21016719; PMID:11130712  
 A/Accession: F86251  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-536 <STO>  
 A/Cross-references: UNIPROT:Q9SA99; GB:AE005172; NID:94835780; PIDN:AAD30246.1; GSPDB:GNC  
 C/Genetics:  
 A/Map position: 1  
 C/Superfamily: poppy reticuline oxidase

Query Match 72.7%; Score 32; DB 2; Length 536;  
 Best Local Similarity 55.6%; Pred. No. 69;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
 DB 286 FLOLMPEPV 294

## RESULT 16

A41120  
 prostaglandin transporter - rat  
 N/Alternate names: matrix F/G  
 C/Species: *Rattus norvegicus* (Norway rat)  
 C/Date: 27-Mar-1992 #sequence\_revision 11-Aug-1995 #text\_change 09-Jul-2004  
 C/Accession: A41120  
 R/Hakes, D.J.; Berzney, R.  
 Proc. Natl. Acad. Sci. U.S.A. 88, 6186-6190, 1991  
 A/Title: Molecular cloning of matrix F/G: a DNA binding protein of the nuclear matrix th  
 A/Reference number: A41120; MUID:91296785; PMID:2068100  
 A/Accession: A41120  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-643 <HAK>  
 A/Cross-references: UNIPROT:Q00910; GB:M64862  
 A/Note: the authors did not translate the codons for residues 1-99 in this reference  
 R/Kana, N.; Lu, R.; Satriano, J.A.; Bao, Y.; Wolkoff, A.W.; Schuster, V.L.  
 Science 268, 866-869, 1995  
 A/Title: Identification and characterization of a prostaglandin transporter.  
 A/Reference number: A38955; MUID:95273959; PMID:7754369  
 A/Contents: annotation  
 C/Keywords: transmembrane protein

Query Match 72.7%; Score 32; DB 2; Length 643;  
 Best Local Similarity 66.7%; Pred. No. 84;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
 DB 318 FLOLMPEPV 326

## RESULT 17

B64126  
 transposase homolog H1478 - *Haemophilus influenzae* (strain Rd KW20)  
 C/Species: *Haemophilus influenzae*



C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004  
 C/Accession: B64126  
 R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
 ; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
 Science 269, 486-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
 A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A/Reference number: A64000; MUID:95350630; PMID:7542800  
 A/Accession: B64126  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A/Residues: 1-687 <TIGR>  
 A/Cross-references: UNIPROT:O05069; GB:U32825; GB:I42023; NID:G3212226; PIDN:AAC23125.1;  
 C/Superfamily: phage Mu transposase

Query Match 72.7%; Score 32; DB 2; Length 687;  
 Best Local Similarity 66.7%; Pred. No. 90;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLQULMBPV 9  
 ||:|||||  
 Db 624 FLEMLMBPV 632

RESULT 18  
 S58108  
 Hypothetical protein SPAC31A2.16 - fission yeast (Schizosaccharomyces pombe)  
 C/Species: Schizosaccharomyces pombe  
 C/Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004  
 C/Accession: J38614; S58108  
 R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, July 1995  
 A/Reference number: Z21731  
 A/Accession: J38614  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A/Residues: 1-1101 <DB2>  
 A/Cross-references: UNIPROT:Q09733; EMBL:Z50113; NID:G914878; PIDN:CAA90474.1; PID:G9148  
 A/Experimental source: strain 972h.; cosmid C31A2  
 C/Genetics:  
 A:Gene: SPAC31A2.16  
 A:Map position: 1  
 A:Insertions: 808/1; 884/1; 920/2  
 C/Superfamily: fission yeast protein SPAC31A2.16; CDC24 homology  
 F:230-431/Domain: CDC24 homology <CD24>

Query Match 72.7%; Score 32; DB 2; Length 1101;  
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 QLMBPV 9  
 ||:|||||  
 Db 379 QLMBPV 385

RESULT 19  
 T16927  
 Hypothetical protein T23F2.2 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C/Accession: T16927  
 R:Du, Z.  
 submitted to the EMBL Data Library, October 1995  
 A/Description: The sequence of C. elegans cosmid T23F2.  
 A/Reference number: Z18608  
 A/Accession: T16927  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A/Residues: 1-1510 <DUZ>  
 A/Cross-references: EMBL:U3649; NID:G1049370; PID:G1049372; PIDN:AAA80384.1; CESP:T23F2  
 C/Genetics:  
 A:Gene: CESP:T23F2.2

A:Insertions: 20/2; 81/1; 119/2; 211/3; 236/3; 290/2; 344/1; 396/3; 439/3; 465/3; 556/3; 60

Query Match 72.7%; Score 32; DB 2; Length 1510;  
 Best Local Similarity 66.7%; Pred. No. 2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLQULMBPV 9  
 ||:|||||  
 Db 429 YLQULMBPV 437

RESULT 20  
 C36179  
 gonadotropin II beta chain precursor - chum salmon  
 C/Species: Oncorhynchus keta (chum salmon)  
 C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: C36179; S09344  
 R:Sekine, S.; Saito, A.; Itoh, H.; Kawachi, H.; Itoh, S.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 8645-8649, 1989  
 A/Title: Molecular cloning and sequence analysis of chum salmon gonadotropin cDNAs.  
 A/Reference number: A36179; MUID:90046849; PMID:2813416  
 A/Accession: C36179  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A/Residues: 1-142 <SEK>  
 A/Cross-references: UNIPROT:P10256; GB:M27154; NID:G213429; PIDN:AAA49409.1; PID:G213430  
 R:Itoh, H.; Suzuki, K.; Kawachi, H.  
 Gen. Comp. Endocrinol. 71, 438-451, 1988  
 A/Title: The complete amino acid sequences of beta-subunits of two distinct chum salmon  
 A/Reference number: S07216; MUID:89053031; PMID:3152067  
 A/Accession: S09344  
 A:Molecule type: protein  
 A/Residues: 24-72; 1, 74-142 <TRO>  
 C/Superfamily: pituitary glycoprotein hormone beta chain  
 C/Keywords: glycoprotein; heterodimer; hormone; pituitary  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-142/Product: gonadotropin II beta chain #status experimental <MAT>  
 F:29-54,43-77,46-108,58-130,92-120,110-113/Disulfide bonds: #status predicted  
 F:33/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 70.5%; Score 31; DB 1; Length 142;  
 Best Local Similarity 66.7%; Pred. No. 29;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLQULMBPV 9  
 ||:|||||  
 Db 13 FLQULMBPV 21

RESULT 21  
 I50143  
 gonadotropin II beta chain - arctic cisco  
 C/Species: Coregonus autumnalis (arctic cisco)  
 C/Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 09-Jul-2004  
 C/Accession: I50143  
 R:Trofimova, I.N.; Belikov, S.I.  
 Mol. Biol. (Mosk.) 28, 1052-1056, 1994  
 A/Title: Cloning and sequencing the cDNA for the beta-subunit of Baikal omul gonadotrop  
 A/Reference number: I50143; MUID:95082790; PMID:7990827  
 A/Accession: I50143  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A/Residues: 1-142 <TRO>  
 A/Cross-references: UNIPROT:P48251; GB:I23431; NID:G387920; PIDN:AAA68207.1; PID:G387921  
 C/Genetics:  
 A:Gene: GTR-II  
 C/Superfamily: pituitary glycoprotein hormone beta chain

Query Match 70.5%; Score 31; DB 2; Length 142;  
 Best Local Similarity 66.7%; Pred. No. 29;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLQULMBPV 9

Db 14 FLQILMEPV 22

## RESULT 22

A69939 conserved hypothetical protein ypmr - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C/Accession: A69939; 140004

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berber C.; Bron, S.; Brouillette, S.; Bruch, C.V.; Caldwell, B.; Capuano, N.M.; Chd A.; Ehrlich, S.D.; Emerson, P.T.; Ertelt, K.D.; Ertelt, C.; Ferraty, E. Nature 390, 249-256, 1997

A/Author: Foulger, D.; Fritz, C.; Fujita, M.; Fuma, S.; Gallazzi, A.; Gallier

leach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lindholm, A.;

Autore: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.;

Autore: Schleich, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serch

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A/Author: Yoshikawa, H.F.; Zumbstein, B.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; NCID:98044033; PMID:9384377

A/Accession: A69939

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-255 <KUN>

A/Cross-references: UNIPROT:P40766; GB:Z99115; GB:AL009126; NID:G2634478; PIDN:CAM14092.

A/Experimental source: strain 168

R/Smith, H.; de Jong, A.; Bron, S.; Venema, G.

Gene 70, 351-361, 1988

A/Title: Characterization of signal-sequence-coding regions selected from the Bacillus s

A/Reference number: I39994; NCID:89108019; PMID:3145906

A/Accession: 140004

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 'M', 4-56, 'OLGARVGYIEIDPLESTAOA' <RES>

A/Cross-references: GB:M22910; NID:G143689; PIDN:AAA2826.1; PID:G143690

C/Genetics:

A/Gene: ypmr

Query Match

Best Local Similarity 70.5%; Score 31; DB 2; Length 255;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQILMEPV 8

Db 130 FLQILMEPV 137

## RESULT 23

T29192 hypothetical protein T03FL.6 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T29192

R/Du, Z.; Le, T.T.

submitted to the EMBL Data Library, February 1997

A/Description: The sequence of C. elegans cosmid T03FL.

A/Reference number: Z20586

A/Accession: T29192

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-342 <DUZ>

A/Cross-references: UNIPROT:P91421; EMBL:U88169; PIDN:AAA42234.1; GSPDB:GNO0019; CESP:TC

C/Genetics:

A/Experimental source: strain Bristol N2; clone T03FL

A/Gene: CESP:T03FL.6

A/Map position: 1

A/Introns: 65/1; 121/1; 147/3; 219/1; 303/3

C/Superfamily: Caenorhabditis elegans hypothetical protein T03FL.6

Query Match

Best Local Similarity 70.5%; Score 31; DB 2; Length 342;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLQILMEPV 9

Db 310 FLQILMEPV 318

## RESULT 24

T01001 hypothetical protein At2g39720 [imported] - Arabidopsis thaliana

N/Alternate names: RING-H2 finger protein RHCA

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004

C/Accession: T01001; T51857; F84820

R/Rounley, S.D.; Lin, X.; Kechum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul,

submitted to the EMBL Data Library, November 1997

A/Description: Arabidopsis thaliana chromosome II BAC T517 genomic sequence.

A/Reference number: Z14162

A/Accession: T01001

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-401 <ROU>

A/Cross-references: UNIPROT:Q22283; EMBL:AC03000; NID:G2642152; PID:G2642154

A/Experimental source: cultivar Columbia

R/Jensen, R.B.; Jensen, K.L.; Jespersen, H.M.; Skriver, K.

FEBS Lett 436, 283-7, 1998

A/Title: Widespread occurrence of a highly conserved RING-H2 zinc finger motif in the mo

A/Reference number: Z13771; NCID:98452956; PMID:9781696

A/Accession: T51857

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-401 <JEN>

A/Cross-references: EMBL:AF079186; PIDN:AA069860.1

R/Lin, X.; Kaul, S.; Rounley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; ;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: AB4420; NCID:20083487; PMID:1061197

A/Accession: F84820

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-401 <STO>

A/Cross-references: GB:AB002093; NID:G2642154; PIDN:AB87121.1; GSPDB:GNO0139

C/Genetics:

A/Gene: T517.2; At2g39720

A/Map position: 2

F/197-247/Domain: RING finger homology <RRN>

Query Match

Best Local Similarity 70.5%; Score 31; DB 2; Length 401;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQILMEPV 9

Db 30 FLQILMEPV 38

## RESULT 25

S69677 hypothetical protein YDR393W - yeast (Saccharomyces cerevisiae)

C/Species: Saccharomyces cerevisiae

C/Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 09-Jul-2004

C/Accession: S69677

R/Dietrich, F.S.

submitted to the EMBL Data Library, July 1995

A/Description: The sequence of S. cerevisiae cosmid 9481, 9509, 9926, 9461, and lambda

A/Reference number: S69677

A/Accession: S69677

A:Molecule type: DNA  
A:Residues: 1-456 <DIE>  
A:Cross-references: UNIPROT:Q04172; EMBL:U32274; NID:g927313; PID:g927326; GSPDB:GN00004  
C/Genetics:  
A:Gene: SGD:SHE9; MIPS:YDR333W  
A:Cross-references: SGD:S0002801  
A:Map position: 4R

Query Match  
Best Local Similarity 70.5%; Score 31; DB 2; Length 456;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FLOLMRPV 8  
Db 313 FLOLMRPV 319

## RESULT 26

T10626  
reticuline oxidase homolog F21C20.190 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C/Accession: T10626  
R:Bayan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke  
submitted to the Protein Sequence Database, June 1999  
A:Reference number: 216991  
A:Accession: T10626  
A:Molecule type: DNA  
A:Residues: 1-539 <BEV>

A:Cross-references: UNIPROT:Q95VG3; EMBL:AL080254; GSPDB:GN00062; ATSP:F21C20.190  
A:Experimental source: cultivar Columbia; BAC clone F21C20  
C/Genetics:  
A:Gene: ATSP:F21C20.190  
A:Map position: 4  
C:Superfamily: poppy reticuline oxidase

Query Match  
Best Local Similarity 70.5%; Score 31; DB 2; Length 539;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLMRPV 9  
Db 292 FLOLMRPV 300

## RESULT 27

A55019  
muscarinic acetylcholine receptor, M3 isoform - chicken  
C/Species: Gallus gallus (chicken)  
C/Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 09-Jul-2004  
C/Accession: A55019  
R:Gaduh, A.P.; Galper, J.B.  
J. Biol. Chem. 269, 25823-25829, 1994  
A:Title: A novel M-3 muscarinic acetylcholine receptor is expressed in chick atrium and  
A:Reference number: A55019; MUID:95014393; PMID:7929287  
A:Accession: A55019  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-639 <GAD>

A:Cross-references: UNIPROT:P49578; GB:L10617; NID:g530097; PID:AAA65961.1; PID:g530098  
A:Superfamily: vertebrate rhodopsin  
C/Keywords: neurotransmitter receptor

Query Match  
Best Local Similarity 70.5%; Score 31; DB 2; Length 639;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLMRPV 9  
Db 270 FLOLMRPV 278

## RESULT 28

T43317  
pgl-1 protein - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C/Accession: T43317  
R:Kawaaki, I.; Shim, Y.H.; Kirchner, J.; Kaminker, J.; Wood, W.B.; Strome, S.  
Cell 94, 635-645, 1998  
A:Title: Pgl-1, a predicted RNA-binding component of germ granules, is essential for fer  
A:Reference number: 222420; MUID:98412659; PMID:9741628  
A:Accession: T43317  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-730 <KAW>

A:Cross-references: UNIPROT:Q9TZQ3; EMBL:AF077868; NID:g3411217; PID:AAC6100.1; PID:g3  
A:Genetics:  
A:Gene: pgl-1  
A:Map position: 4

Query Match  
Best Local Similarity 70.5%; Score 31; DB 2; Length 730;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMRPV 9  
Db 220 FLOLMRPV 228

## RESULT 29

T29177  
hypothetical protein ZK381.4 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C/Accession: T29177  
R:Latreille, P.; Stelljes, L.  
submitted to the EMBL Data Library, March 1996  
A:Description: The sequence of C. elegans cosmid ZK381.  
A:Reference number: 220583  
A:Accession: T29177  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-771 <LAT>

A:Cross-references: EMBL:U52003; PID:AAA96096.1; GSPDB:GN00022; CESP:ZK381.4  
A:Experimental source: strain Bristol N2; clone ZK381  
C/Genetics:  
A:Gene: CESP:ZK381.4  
A:Map position: 4  
A:Introns: 36/3; 80/3; 129/3; 249/2; 363/1; 410/3; 468/1; 642/3

Query Match  
Best Local Similarity 70.5%; Score 31; DB 2; Length 771;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMRPV 9  
Db 261 FLOLMRPV 269

## RESULT 30

F81294  
probable helicase Cj1481c [imported] - Campylobacter jejuni (strain NCTC 11168)  
C/Species: Campylobacter jejuni  
C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C/Accession: F81294  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrer  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: F81294  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-921 <PAR>

A:Cross-references: UNIPROT:Q9PM11; GB:AL139078; GB:AL111168; NID:g6968723; PID:CAB7390.

A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj1481c

Query Match  
Best Local Similarity 70.5%; Score 31; DB 2; Length 921;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLOLMPEV 8  
|||:|  
Db 597 FLOLMPEP 604

RESULT 31

D59435  
Gem-interacting protein [imported] - human

C:Species: Homo sapiens (man)  
C>Date: 03-Jun-2002 #sequence\_revision 03-Jun-2002 #text\_change 09-Jul-2004

C:Accession: D59435

R:Arresta, S.; Berger, F.; Berger, R.; de Gunzburg, J.

submitted to GenBank, November 2000

A:Description: GMIP, a Gem interacting protein.

A:Reference number: D59435

A:Accession: D59435

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-970 <ARE>

A:Cross-references: UNIPROT:Q9P107; GB:NP\_057657; PID:g7706107; PIDN:NP\_057657.1

Query Match  
Best Local Similarity 70.5%; Score 31; DB 2; Length 970;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 FLOLMPEV 9  
|||:|  
Db 626 FLOLMPEP 634

RESULT 32

A35098  
MHC class III histocompatibility antigen HLA-B-associated transcript 3 - human

C:Species: Homo sapiens (man)  
C>Date: 10-Aug-1990 #sequence\_revision 10-Aug-1990 #text\_change 09-Jul-2004

C:Accession: A35098

R:Bannerji, J.; Sands, J.; Strominger, J.L.; Spies, T.

Proc. Natl. Acad. Sci. U.S.A. 87, 2374-2378, 1990

A>Title: A gene pair from the human major histocompatibility complex encodes large prot

A:Reference number: A35098; MUID:90192810; PMID:2156268

A:Accession: A35098

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1132 <BAN>

A:Cross-references: UNIPROT:P46379; GB:M3519; NID:G179346; PIDN:AA35587.1; PID:G179347

C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

F:17-91/Domain: ubiquitin homology <UBH>

Query Match  
Best Local Similarity 70.5%; Score 31; DB 2; Length 1132;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 QOLMPEV 9  
|||:|  
Db 547 QOLMPEV 553

RESULT 33

E86201

protein F12K11.4 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: E86201

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso, J.

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

aneen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani, L.

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shum, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86201

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-11760 <STO>

A:Cross-references: UNIPROT:Q9SHK6; GB:AE005172; NID:g6692694; PIDN:AA24828.1; GSPDB:GNC

C:Genetics:

A:Gene: F12K11.4

A:Map position: 1

Query Match  
Best Local Similarity 70.5%; Score 31; DB 2; Length 1760;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLOLMPEV 8  
|||:|  
Db 927 FLOLMPEP 934

RESULT 34

E83533  
hypothetical protein PA0900 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004

C:Accession: E83533

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lapid, K.; Lim,

Lo, S.; Olson, M.V.

Nature 406, 959-964, 2000

A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: E83533

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-96 <STO>

A:Cross-references: UNIPROT:O50176; GB:AE004524; GB:AE004091; NID:g9946795; PIDN:AA60428

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0900

Query Match  
Best Local Similarity 68.2%; Score 30; DB 2; Length 96;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LQLMPEV 9  
|||:|  
Db 71 LQLMPEP 78

RESULT 35

S56917

hypothetical protein YUL135w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein Y0666

C:Species: Saccharomyces cerevisiae

C>Date: 08-Jul-1995 #sequence\_revision 08-Sep-1995 #text\_change 09-Jul-2004

C:Accession: S56917; S71662

R:Katsoulou, C.; Tzermia, M.; Alexandraki, D.

submitted to the Protein Sequence Database, September 1995

A:Reference number: S56912

A:Accession: S56917

A:Molecule type: DNA

A:Residues: 1-105 <RXT>

A:Cross-references: UNIPROT:P47012; EMBL:Z49410; NID:G1008339; PIDN:CAA89429.1; PID:G100

R:Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.

Yeast 12, 787-797, 1996

A:Title: Sequence analysis of a 40.7 kb segment from the left arm of yeast chromosome X of chromosome XI.  
A:Reference number: S71643; MUID:96408771; PMID:8813765  
A:Accession: S71662  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-105 <KAM>  
A:Cross-references: EMBL:X87371; NID:G854542; PIDN:CA60820.1; PID:G992620  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1995  
C:Genetics:  
A:Cross-references: SGD:S0003671  
A:Map position: 10L  
C:Superfamily: Saccharomyces hypothetical protein YOL135W

Query Match 68.2%; Score 30; DB 2; Length 105;  
Best Local Similarity 75.0%; Pred. No. 35;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEV 8  
|:|:|:|  
Db 2 FNOILLBP 9

RESULT 36  
S45797  
probable membrane protein YBL062W - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YBL0505  
C:Species: Saccharomyces cerevisiae  
C>Date: 26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Jul-2004  
R:Dubois, B.; El Bakoury, M.; Glansdorff, N.; Mesenguy, F.; Pierard, A.; Scherens, B.; submitted to the Protein Sequence Database, August 1994  
A:Reference number: S45782  
A:Accession: S45797  
A:Molecule type: DNA  
A:Residues: 1-126 <DUB>  
A:Cross-references: UNIPROT:P38189; EMBL:Z35823; NID:G536095; PID:G536096; GSPDB:GN00002  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: MIPS:YBL062W  
A:Cross-references: SGD:S0000158  
A:Map position: 2L  
C:Keywords: transmembrane protein  
F:8-28/Domain: transmembrane #status predicted <TM>  
F:71-99/Domain: transmembrane #status predicted <TM2>

Query Match 68.2%; Score 30; DB 2; Length 126;  
Best Local Similarity 66.7%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLMPEV 9  
|:|:|:|  
Db 73 FLMVSEPV 81

RESULT 37  
A25800  
gonadotropin beta chain precursor - chinook salmon  
C:Species: Oncorhynchus tshawytscha (chinook salmon)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: A25800  
R:Trinh, K.Y.; Wang, N.C.; Hew, C.L.; Crim, L.W.  
Eur. J. Biochem. 159, 619-624, 1986  
A:Title: Molecular cloning and sequencing of salmon gonadotropin beta subunit.  
A:Reference number: A25800; MUID:87004682; PMID:2428617  
A:Accession: A25800  
A:Molecule type: mRNA  
A:Residues: 1-142 <TRI>  
A:Cross-references: UNIPROT:P07732; GB:X04404; NID:G64202; PIDN:CA27992.1; PID:G64203  
C:Superfamily: pituitary glycoprotein hormone beta chain  
F:29-54,43-77,46-108,58-130,92-120,110-113/Disulfide bonds: #status predicted

Query Match 68.2%; Score 30; DB 1; Length 142;

Best Local Similarity 55.6%; Pred. No. 47;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEV 9  
|:|:|:|  
Db 13 FLOILLRPI 21

RESULT 38  
T29222  
hypothetical protein F55G1.7 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T29222  
R:Murray, J.; Le, T.T.  
submitted to the EMBL Data Library, May 1996  
A:Description: The sequence of C. elegans cosmid F55G1.  
A:Reference number: Z20591  
A:Accession: T29222  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-307 <MGR>  
A:Cross-references: UNIPROT:Q20844; EMBL:U58750; PIDN:AAB0641.1; GSPDB:GN00022; CESP:F5  
A:Experimental source: strain Bristol N2; clone F55G1  
C:Genetics:  
A:Gene: CESP:F55G1.7  
A:Map position: 4  
A:Insertions: 64/2; 160/2; 193/2; 250/2  
C:Superfamily: Caenorhabditis elegans hypothetical protein F55G1.7

Query Match 68.2%; Score 30; DB 2; Length 307;  
Best Local Similarity 75.0%; Pred. No. 1,1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEV 8  
|:|:|:|  
Db 21 FLOLMPEV 28

RESULT 39  
G75117  
dipeptide abc transporter, ATP-binding protein (dppf) PAB1875 - Pyrococcus abyssi (strain dipeptide abc transporter, ATP-binding protein (dppf) PAB1875 - Pyrococcus abyssi (strain  
C:Species: Pyrococcus abyssi  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: G75117  
R:anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc  
A:Reference number: A75001  
A:Accession: G75117  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-326 <KAM>  
A:Cross-references: UNIPROT:Q9V0P7; GB:A248285; GB:AL096836; NID:G5458067; PIDN:CAB4965  
A:Experimental source: strain Orsay  
C:Genetics:  
A:Gene: PAB1875  
C:Superfamily: inner membrane protein malK; ATP-binding cassette homology  
F:30-222/Domain: ATP-binding cassette homology <ABC>

Query Match 68.2%; Score 30; DB 2; Length 326;  
Best Local Similarity 71.4%; Pred. No. 1,1e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 QLMPEV 9  
|:|:|:|  
Db 112 QLMPEV 118

RESULT 40  
G71014  
probable dipeptide transport ATP-binding protein dppf - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii

C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 16-Aug-2004  
C;Accession: G71014  
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A;Reference number: A71000; MUID:98344137; PMID:9679194  
A;Accession: G71014  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-329 <KAW>  
A;Cross-references: UNIPROT:O50121; GB:AP000006; NID:g3236133; PIDN:BAA30519.1; PID:g325  
A;Experimental source: strain OT3  
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C;Genetics:  
A;Note: PH1413  
C;Function:  
A;Description: probably responsible for energy-coupling to the transport system  
C;Keywords: ATP; binding protein-dependent transport system; membrane protein; nucleotid  
F;33-232/Domain: ATP-binding cassette homology <ABC>  
F;50-57/Region: nucleotide-binding motif A (P-loop)  
F;176-180/Region: nucleotide-binding motif B

Query Match 68.2%; Score 30; DB 2; Length 329;  
Best Local Similarity 71.4%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 QULMEPV 9  
|:|||||  
DB 115 QULMEPV 121

Search completed: January 12, 2005, 20:15:41  
Job time : 26.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 12, 2005, 19:38:08 ; Search time 85 Seconds  
(without alignments)  
60.922 Million cell updates/sec

Title: US-09-870-216C-3  
Perfect score: 44  
Sequence: 1 FLOLMEPV 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: UniProt\_02:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	37	84.1	75	2	Q8PYR9
2	35	79.5	47	2	Q85X02
3	35	79.5	329	2	Q8SQ02
4	35	79.5	379	2	003991
5	34	77.3	297	1	BNHL_CAMEL
6	34	77.3	318	2	Q8DN74
7	34	77.3	318	2	Q8ES57
8	34	77.3	318	2	Q97NY1
9	34	77.3	471	2	Q8AVC6
10	34	77.3	760	2	Q9NIH0
11	34	77.3	809	2	Q8S159
12	34	77.3	1083	2	Q7SBB3
13	34	77.3	1869	2	Q7QUP7
14	34	77.3	2785	2	Q75691
15	33	75.0	113	2	Q9SA20
16	33	75.0	163	2	Q85FQ9
17	33	75.0	226	2	Q87826
18	33	75.0	252	2	Q6IRDS
19	33	75.0	252	2	AAH70960
20	33	75.0	287	2	Q7NGE4
21	33	75.0	295	2	Q7NGE4
22	33	75.0	318	2	Q6G896
23	33	75.0	318	2	Q6G107
24	33	75.0	318	2	Q8NKM9
25	33	75.0	318	2	Q99VP0
26	33	75.0	318	2	Q7A669
27	33	75.0	319	2	Q8CPZ9
28	33	75.0	337	2	Q7ZSK8
29	33	75.0	366	2	Q84132
30	33	75.0	380	2	Q6TLD0
31	33	75.0	380	2	AAQ94603

32	33	75.0	382	2	Q6PHE9	Q6PHE9 brachydanio
33	33	75.0	382	2	AAH56578	AAH56578 brachydan
34	33	75.0	409	1	R23B_HUMAN	P54727 homo sapien
35	33	75.0	415	2	Q6NVC3	Q6NVC3 mus musculu
36	33	75.0	415	2	AAH68193	AAH68193 mus muscu
37	33	75.0	416	1	R23B_MOUSE	P54728 mus musculu
38	33	75.0	435	2	Q6FV70	Q6FV70 candia gla
39	33	75.0	1569	2	Q8S012	Q8S012 oryza sativ
40	33	75.0	2746	2	Q8IMS2	Q8IMS2 drosophila
41	33	75.0	2746	2	AAH14067	AAH14067 drosophila
42	33	75.0	2764	2	001399	001399 drosophila
43	33	75.0	2802	2	001397	001397 drosophila
44	33	75.0	2802	2	001398	001398 drosophila
45	33	75.0	2802	2	Q9VB02	Q9VB02 drosophila

## ALIGNMENTS

```

RESULT 1
Q8PYR9 PRELIMINARY; PRT; 75 AA.
ID Q8PYR9
AC Q8PYR9;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein MM0791.
GN OrderedLocName=MM0791;
OS Methanosarcina mazel (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanocicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / Gol / ATCC BAA-199 / DSM 3647 / OCM 88;
RA MEDLINE=2120827; PubMed=1215824;
RA Deppenmeier U., Johann A., Hartoch T., Merk R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wietzer A., Baumer S., Jacob C.,
RA Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,
RA Bhattacharya A., Lykdis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazel: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461 (2002).
KW EMBL; AEO13304; AAM30487.1; -.
SQ Complete proteome; Hypothetical protein.
SQ SEQUENCE 75 AA; 9035 MW; 695543A5C06F505D CRC64;

Query Match 84.1%; Score 37; DB 2; Length 75;
Best Local Similarity 87.5%; Pred. No. 4.5;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMEPV 8
Db 40 FLOLMEPV 47

RESULT 2
Q85X02 PRELIMINARY; PRT; 47 AA.
ID Q85X02;
AC Q85X02;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE ORF47d.
OS Pinus koraiensis (Korean pine).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=88728;
RN [1]
RP SEQUENCE FROM N.A.
RA Noh E.W., Lee J.S., Choi Y.I., Han M.S., Yi Y.S., Han S.U.;

```

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY228468; AA074065.1; -  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 KM Chloroplast.  
 SQ SEQUENCE 47 AA; 5510 MW; B72C7950F3876DCD CRC64;

Query Match 79.5%; Score 35; DB 2; Length 47;  
 Best Local Similarity 87.5%; Pred. No. 7.7;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMPEV 8  
 Db 25 FLOLMPEV 32

RESULT 3  
 ID 08SQ02 PRELIMINARY; PRT; 329 AA.

AC 08SQ02;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)  
 DE 01-MAR-2004 (TREMBLrel. 26, last annotation update)  
 DE Pregnane X receptor (Fragment).  
 GN Name=PXK;  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 ON NCBI\_TaxId=9615;  
 RX [1]

RP SEQUENCE FROM N.A.  
 RA MEDLINE=2197871; PubMed=11981033;  
 RA Moore L.B., Maglich J.M., McKee D.D., Wisely B., Willson T.M.,  
 RA Kiewer S.A., Lambert M.H., Moore J.T.;  
 RT "Pregnane X receptor (PXK), constitutive androstane receptor (CAR),  
 RT and benzocate X receptor (BXR) define three pharmacologically distinct  
 RT classes of nuclear receptors";  
 RL Mol. Endocrinol. 16:977-986(2002).

DR EMBL; AF454670; AAM10632.1; -  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR000536; Hmon\_recept\_11g.  
 DR InterPro; IPR0011723; Stchrnm\_receptor.  
 DR InterPro; IPR008946; Str\_ncl\_receptor.  
 DR InterPro; IPR000324; Vld\_ncl\_receptor.  
 DR Pfam; PF00104; Hormone\_recep. 1.  
 DR PRINTS; PR00398; STRDHOMONER.  
 DR PRINTS; PR00350; VITAMINDR.  
 DR SMART; SM00430; HOL1; 1.  
 DR KX Receptor.  
 FT NON TER

SQ SEQUENCE 329 AA; 37479 MW; 237F59C90C67474 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 329;  
 Best Local Similarity 77.8%; Pred. No. 59;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMPEV 9  
 Db 210 FLOLMPEV 218

RESULT 4  
 ID 003991 PRELIMINARY; PRT; 379 AA.

AC 003991;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)  
 DE 01-MAR-2004 (TREMBLrel. 26, last annotation update)  
 DE RAD23 protein, isoform II.  
 OS Daucus carota (Carrot).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Campanulids; Apiales; Apiaceae; Apioidae; Scandiceae; Daucinae;  
 OC Daucus.  
 ON NCBI\_TaxId=4039;  
 RX [1]

RP SEQUENCE FROM N.A.  
 RA MEDLINE=98345997; PubMed=9681019;  
 RA Sturm A., Leinhardt S.;  
 RT "Two isoforms of plant RAD23 complement a UV-sensitive rad23 mutant in  
 RT yeast";  
 RL Plant J. 13:815-821(1998).

DR EMBL; Y12014; CAAT7242.1; -  
 DR PIR; T14337; T14337.  
 DR HSSP; P54725; IF4I.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0006289; P:nucleotide-excision repair; IEA.  
 DR InterPro; IPR004806; Rad23.  
 DR InterPro; IPR006636; STI1.  
 DR InterPro; IPR000449; UBA.  
 DR InterPro; IPR009060; UBA\_1like.  
 DR InterPro; IPR000626; Ubiquitin.  
 DR Pfam; PF00627; UBA; 2.  
 DR Pfam; PF00240; ubiquitin; 1.  
 DR PRINTS; PR01839; RAD23PROTEIN.  
 DR SMART; SM00727; STI1; 1.  
 DR SMART; SM00165; UBA; 2.  
 DR SMART; SM00213; UBQ; 1.  
 DR TIGRFAMs; TIGR00601; rad23; 1.  
 DR PROSITE; PS50030; UBA; 2.  
 DR PROSITE; PS50053; UBQUITIN 2; 1.

SQ SEQUENCE 379 AA; 40530 MW; B26697B439C5929 CRC64;

Query Match 79.5%; Score 35; DB 2; Length 379;  
 Best Local Similarity 77.8%; Pred. No. 68;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMPEV 9  
 Db 302 FLOLMPEV 310

RESULT 5  
 ID RNHL\_CABEL STANDARD; PRT; 297 AA.

AC 0906P;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, last sequence update)  
 DT 05-JUL-2004 (Rel. 44, last annotation update)  
 DE Ribonuclease HI large subunit (EC 3.1.26.-) (RNase HI large subunit)  
 DE (RNase H(35)).  
 GN Name=rnh-2; Synonyms=rnh2; ORNames=TI3H5.7;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 ON NCBI\_TaxId=6239;  
 RX [1]

RP SEQUENCE FROM N.A.  
 RA MEDLINE=22300325; PubMed=12411600;  
 RA Arundhandran A.P., Cerritelli S.M., Bowen N.J., Chen X., Krause M.W.,  
 RA Crouch R.J.;  
 RT "Multiple ribonucleases H-encoding genes in the Caenorhabditis elegans  
 RT genome contrasts with the two typical ribonuclease H-encoding genes in  
 RT the human genome";  
 RL Mol. Biol. Evol. 19:1910-1919(2002).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=Briscot N2;  
 RA lightning J.;  
 RA Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 RL -1- FUNCTION: Degrades the ribonucleotide moiety on RNA-DNA hybrid  
 CC molecules. Participates in DNA replication (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomonoester.



CC -1- SIMILARITY: Belongs to the RNase HII family, Eukaryotic subfamily.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; AF181619; AAF01208.1; -  
 DR EMBL; Z66524; CAC70103.1; -  
 DR HSPR; Q57599.1; EXE.  
 DR Wormpep; T13H5.7; CE28960.  
 DR InterPro; IPR01352; RNase\_HII/HIII.  
 DR InterPro; IPR004649; RnhII.  
 DR Pfam; PF01351; RNase\_HII; 1.  
 DR TIGRFAMs; TIGR00729; RnhII; 1.  
 KM Endonuclease; Hydrolyase; Nuclease.  
 FT ACT\_SITE 27 27 By similarity.  
 FT ACT\_SITE 138 138 By similarity.  
 FT ACT\_SITE 166 166 By similarity.  
 SQ SEQUENCE 297 AA; 33191 MW; 233C11EDD4A7B5B2 CRC64;  
 Query Match 77.3%; Score 34; DB 1; Length 297;  
 Best Local Similarity 77.8%; Pred. No. 87;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 FLOQLMEPV 9  
 Db 217 FLOQLMEPV 225  
 RESULT 6  
 Q8DNM4 PRELIMINARY; PRT; 318 AA.  
 ID Q8DNM4  
 AC Q8DNM4;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE ABC transporter membrane-spanning permease-ferric iron transport.  
 GN Name-fact; OrderedlocusNames=sp1685;  
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 NC NCB1\_TaxID=171101;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC BAA-255 / R6;  
 RX MEDLINE=21429245; PubMed=11544234;  
 RA Hoskins J., Alborn W.E., Jr., Arnold J., Blaszcak L.C., Burgett S.,  
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,  
 RA Gilmour R., Glas J.S., Khoja H., Kraft A.R., Lagace R.E.,  
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushita P.,  
 RA McAnen J.M., McHenry M., Mcleaster K., Mundy C.W., Nicas T.I.,  
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,  
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,  
 RA Zook C.A., Baltz R.H., Jaeknas S.R., Rostock P.R., Jr., Skatrud P.L.,  
 RA Glas J.I.;  
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6."  
 RL J. Bacteriol. 183:5709-5717(2001).  
 DR EMBL; AE008534; AAL00468.1; -  
 DR PIR; C98082; C98082.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR00522; FeCD; 1.  
 DR Pfam; PF01032; FeCD; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 318 AA; 36296 MW; 9C1894E4F57C3210 CRC64;  
 Query Match 77.3%; Score 34; DB 2; Length 318;  
 Best Local Similarity 75.0%; Pred. No. 94;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 FLOQLMEPV 8  
 Db 152 FLOQLMEPV 159  
 RESULT 7  
 Q8ES87 PRELIMINARY; PRT; 318 AA.  
 ID Q8ES87  
 AC Q8ES87;  
 DT 01-MAR-2003 (TReMBLrel. 23, Created)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Ferrichrome ABC transporter permease.  
 GN OrderedlocusNames=OB0540;  
 OS Oceanobacillus thelyensis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
 NC NCB1\_TaxID=182710;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HTE831;  
 RX MEDLINE=2220767; PubMed=12235376;  
 RA Takami H., Takaki Y., Uchiyama I.;  
 RT Ridge and its unexpected adaptive capabilities to extreme  
 RT environments."  
 RL Nucleic Acids Res. 30:3927-3935(2002).  
 DR EMBL; AP004594; BAC12496.1; -  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR00522; FeCD.  
 DR Pfam; PF01032; FeCD; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 318 AA; 35775 MW; 3A53578580A52E71 CRC64;  
 Query Match 77.3%; Score 34; DB 2; Length 318;  
 Best Local Similarity 66.7%; Pred. No. 94;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 FLOQLMEPV 9  
 Db 150 FLOQLMEPV 158  
 RESULT 8  
 Q97NY1 PRELIMINARY; PRT; 318 AA.  
 ID Q97NY1  
 AC Q97NY1;  
 DT 01-OCT-2001 (TReMBLrel. 18, Created)  
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Iron-compound ABC transporter, permease protein.  
 GN OrderedlocusNames=SP1870;  
 OS Streptococcus pneumoniae.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 NC NCB1\_TaxID=1313;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC BAA-334 / TIGR4;  
 RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;  
 RA Tetteil H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
 RA Peterson S.N., Heidelberg J.F., Deboy R.T., Haft D.H., Dodson R.J.,  
 RA Durkin A.S., Gwinn W.L., Kolonay J.F., Nelson W.C., Peterson J.D.,  
 RA Unayam L.A., White O., Salzberg S.L., Lewis M.R., Raddue D.,  
 RA Holtzapple E.K., Knouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,  
 RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,  
 RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
 RT "Complete genome sequence of a virulent isolate of Streptococcus  
 pneumoniae.";

RL Science 293:498-506(2001).  
DR EMBL; AE007478; AAK75942.1; -  
DR PIR; E95218; E95218.  
DR TIGR; SP1870; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000522; FeCD.  
DR Pfam; PF01032; FeCD; 1.  
DR Complete proteome.  
KW SEQUENCE 318 AA; 36234 MW; 72BB0DC3DC0BFD13 CRC64;  
SQ  
Query Match 77.3%; Score 34; DB 2; Length 318;  
Best Local Similarity 75.0%; Pred. No. 94;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLOLMERP 8  
DB 152 FLOVLMRP 159  
RESULT 9  
ID Q9AVC6 PRELIMINARY; PRT; 471 AA.  
AC Q9AVC6;  
DT 01-MAR-2003 (TReMBLrel. 23, Created)  
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE MG52635 protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
OC Xenopodinae; Xenopus.  
NCBI\_TaxID=8355;  
OX  
RN [1] SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RC MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schlier G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gamaralle P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalton D.E., Schmech A., Schein J.B.,  
RA Jones S.J., Maira M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2] SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RC MEDLINE=22341132; PubMed=12454917;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative";  
RL Dev. Dyn. 225:384-391(2002).  
RN [3] SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RA Klein S., Strausberg R.,  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC004171; AAH4171.1; -

DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR PROSITE; PS00850; MFS; 1.  
KW SEQUENCE 471 AA; 51058 MW; 1504C5B9AA70B01 CRC64;  
SQ  
Query Match 77.3%; Score 34; DB 2; Length 471;  
Best Local Similarity 66.7%; Pred. No. 1,46+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLOLMERP 9  
DB 121 FLOVLMRP 129  
RESULT 10  
ID Q9N1H0 PRELIMINARY; PRT; 760 AA.  
AC Q9N1H0;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
DE Elongation factor 2 (Fragment).  
GN Name=EF2;  
OS Stylonychia mytilus.  
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;  
OC Stichotrichida; Oxytrichidae; Stylonychia.  
NCBI\_TaxID=5950;  
OX  
RN [1] SEQUENCE FROM N.A.  
RP MEDLINE=20269356; PubMed=10811219;  
RA Moreira D., Le Guyader H., Philippe H.;  
RT "The origin of red algae and the evolution of chloroplasts";  
RL Nature 405:69-72(2000).  
DR EMBL; AF213664; AAF71707.1; -  
DR GO; GO:0005525; P:GTP binding; IEA.  
DR GO; GO:0003746; P:translation elongation factor activity; IEA.  
DR GO; GO:0006412; P:protein biosynthesis; IEA.  
DR GO; GO:0006414; P:translational elongation; IEA.  
DR InterPro; IPR000640; EFG\_C.  
DR InterPro; IPR009022; EFG\_TII\_V.  
DR InterPro; IPR005517; EFG\_IV.  
DR InterPro; IPR004161; EFTU\_D2.  
DR InterPro; IPR000795; ProctSyn\_GRPbind.  
DR InterPro; IPR005225; Small\_GTP.  
DR InterPro; IPR009000; Translat\_factor.  
DR Pfam; PF00679; EFG\_C; 1.  
DR Pfam; PF03764; EFG\_IV; 1.  
DR Pfam; PF00009; GTP\_EFTU; 1.  
DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
DR PRINTS; PRO0315; ELONGATINFACT.  
DR TIGRFAIR; TIGR00231; small\_GTP; 1.  
DR PROSITE; PS00301; EFACTOR\_GTP; 1.  
KW Elongation factor; GTP-binding; Protein biosynthesis.  
FT NON TER 1  
FT NON TER 1  
FT NON TER 1  
SQ SEQUENCE 760 AA; 84711 MW; 6CD14EB3F1E6590D CRC64;  
Query Match 77.3%; Score 34; DB 2; Length 760;  
Best Local Similarity 55.6%; Pred. No. 2,3e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLOLMERP 9  
DB 237 FVOFIMEPI 245  
RESULT 11  
ID Q8S159 PRELIMINARY; PRT; 809 AA.  
AC Q8S159;  
DT 01-JUN-2002 (TReMBLrel. 21, Created)

RN NCBI\_TaxId=5141;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OR74A;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
 RA Klitting T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Ooi D., Ianklev P., Pedersen D., Nelson M., Washburne M.,  
 RA Seltehrnikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
 RA Kothe G.O., Jedd G., Wewes W., Staben C., Marcotte E., Greenberg D.,  
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
 RA Kamal M., Kanvyessels M., Mauceli E., Bielke C., Rudd S., Frisman D.,  
 RA Kryukova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,  
 RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,  
 RA Deonou C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
 RA Yarden O., Plamann M., Sella S., Dunlap J., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbold D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nisbaum C., Birren B.;  
 RA "The Genome Sequence of the Filamentous Fungus *Neurospora crassa*.";  
 RL Nature 0:0-0(2003).  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL; AA001000173; EAA33646.1; -;  
 DR GO; GO:0005525; F:GTP binding; IEA.  
 DR GO; GO:0003746; F:translation elongation factor activity; IEA.  
 DR GO; GO:0006412; P:protein biosynthesis; IEA.  
 DR GO; GO:0006414; P:translational elongation; IEA.  
 DR InterPro; IPR000640; EFG\_C.  
 DR InterPro; IPR005517; EFG\_IV.  
 DR InterPro; IPR004161; EFTU\_D2.  
 DR InterPro; IPR000795; Protsyn\_GTPbind.  
 DR InterPro; IPR005225; Small\_GTP.  
 DR Pfam; PF00679; EFG\_C; 1.  
 DR Pfam; PF03764; EFG\_IV; 1.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR PRINTS; PR00315; EFONGAKINFC.  
 DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
 DR GTP-binding; Hypothetical protein; Protein biosynthesis.  
 KW KW  
 SQ SEQUENCE 1083 AA; 119199 MW; A2CD42DD89C64731 CRC64;  
  
 Query Match 77.3%; Score 34; DB 2; Length 1083;  
 Best Local Similarity 55.6%; Pred. No. 3.4e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 1 FLOUTIMEPV 9  
 Db 312 FVQGLVEPI 320  
  
 RESULT 13  
 Q7QUP7 PRELIMINARY; PRT; 1869 AA.  
 AC Q7QUP7;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)  
 DE GUP 47 16110 21719.  
 OS *Giardia lamblia* ATCC 50803.  
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
 OX NCBI\_TaxID=184922;  
 RX NCI  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WB C6;  
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,  
 RA Olsen G.U., Sogin M.L.;  
 RA "Draft sequence of the *Giardia lamblia* genome.";  
 RL Submitted (MAR-2003) to the EMBL/Genbank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL; AAC001000091; EAA38765.1; -.

SQ SEQUENCE 1869 AA; 209537 MW; 55EE7149EDA47640 CRC64;  
 Query Match 77.3%; Score 34; DB 2; Length 1869;  
 Best Local Similarity 75.0%; Pred. No. 6e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOQLMEP 8  
 Db 1481 FLELLLEP 1488

RESULT 14  
 075691 PRELIMINARY; PRT; 2785 AA.  
 AC 075691;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE DRIM protein.  
 GN Name=drim;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP MEDLINE=98338066; PubMed=9673349;  
 RA Schirke M., Gnikke A., Bork P., Tarin D., Weidle U.H.;  
 RT "Differential gene expression in mammary carcinoma cell lines";  
 RL Anticancer Res. 18:1409-1421(1998).  
 DR EMBL; AJ006778; CA07243.1; -.  
 DR SWISS-2DPAGE; 075691; -.  
 DR GO; GO:0008285; Pnegative regulation of cell proliferation; TAS.  
 DR InterPro; IPR008938; ARM.  
 DR Pfam; PF07539; DRIM.  
 DR SEQUENCE 2785 AA; 318423 MW; 6A2BBE9E82ADB983 CRC64;

Query Match 77.3%; Score 34; DB 2; Length 2785;  
 Best Local Similarity 77.8%; Pred. No. 9.1e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOQLMEP 9  
 Db 1056 FLDLLEP 1064

RESULT 15  
 09SA20 PRELIMINARY; PRT; 113 AA.  
 AC 09SA20;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
 DE F309.1 protein.  
 GN Name=F309.1;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 CX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vysotskaia V.S., Schwartz J., Yu G., Toriumi M., Lenz C., Liu S.,  
 Lee J., Li J., Kremetskaia I., Liu A., Luros J., Gonzalez A.,  
 Altafi H., Arzajo R., Chao Q., Conn L., Conway A.B., Dunn P.,  
 Hansen N., Hultar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,  
 Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Theologis;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC006341; AAD34676.1; -.  
 DR PIR; H86296; H86296.  
 DR HSSP; P54725; 1F4I.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0006289; P:nucleotide-excision repair; IEA.  
 DR InterPro; IPR009020; Prot\_inh\_propept.  
 DR InterPro; IPR004806; Rad23.  
 DR InterPro; IPR004449; UBA.  
 DR InterPro; IPR009060; UBA\_like.  
 DR Pfam; PF00627; UBA.1.  
 DR PRINTS; PRO1819; RAD23PROTEIN.  
 DR SMART; SM00165; UBA.1.  
 DR PROSITE; PSS0030; UBA.1.  
 DR SEQUENCE 113 AA; 12947 MW; 8BE7CB15C9E12380 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 113;  
 Best Local Similarity 87.5%; Pred. No. 53;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOQLMEP 8  
 Db 34 FLOQLMEP 41

RESULT 16  
 085FQ9 PRELIMINARY; PRT; 163 AA.  
 AC 085FQ9;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE ATP synthase CF1 delta chain.  
 GN Name=atpd;  
 OS Cyanidioschyzon merolae (Red alga).  
 CC Chloroplast.  
 CC Eukaryota; Rhodophyta; Bangiophyceae; Cyanidiales; Cyanidiaceae;  
 CC Cyanidioschyzon.  
 CX NCBI\_TaxID=45157;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=10D;  
 RX MEDLINE=22639682; PubMed=12755171;  
 RA Ohta N., Matsuzaki M., Miumi O., Miyagishima S., Nozaki H.,  
 Tanaka K., Shin-I T., Kohara Y., Kuroiwa T.;  
 RT "Complete Sequence and Analysis of the Plastid Genome of the  
 Unicellular Red Alga Cyanidioschyzon merolae";  
 RL DNA Res. 10:67-77(2003).  
 CC -1- FUNCTION: This protein seems to be part of the stalk that links  
 CF(0) to CF(1). It either transmits conformational changes from  
 CF(0) into CF(1) or is implicated in proton conduction (by  
 similarity).  
 CC -1- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic  
 core - and CF(0) - the membrane proton channel. CF(1) has five  
 subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)  
 has three main subunits: a, b and c (by similarity).  
 CC -1- SIMILARITY: Belongs to the ATPase delta chain family.  
 DR EMBL; AB002583; BAC76286.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.  
 DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.  
 DR GO; GO:0046961; F:hydrogen-transporting ATPase activity; IEA.  
 DR GO; GO:0016787; F:hydrogen-transporting ATPase activity; IEA.  
 DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.  
 DR InterPro; IPR00711; ATPynt\_OSCP.  
 DR Pfam; PF00213; OSCP.1.  
 DR PRINTS; PR00125; ATPASEDELTA.  
 DR TIGRFAms; TIGR01145; ATP\_synth\_delta.1.  
 DR ATP synthetase; CF(1); Chloroplast; Hydrolyase.  
 KW Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 SO SEQUENCE 163 AA; 19233 MW; 448F2F9548FC51D1 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 163;  
 Best Local Similarity 75.0%; Pred. No. 77;



RP SEQUENCE FROM N.A.  
RC Tissue-Lung;  
RA Strauberg R.;  
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
RW EMBL, BC070960, AA070960.1; --  
KW Hypothetical protein.  
FT NON TER  
SQ SEQUENCE 252 AA; 2669 MW; F68AC957DD42445E CRC64;

Query Match 75.0%; Score 33; DB 2; Length 252;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMPEV 9  
Db 161 FLOLMPEV 169

RESULT 20  
Q87GX1 PRELIMINARY; PRT; 287 AA.  
AC 087GX1  
DT 01-JUN-2003 (TRENBLrel. 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE Hypothetical protein VPA1194.  
GN OrderedLocustNames=VPA1194;  
OS Vibrio parahaemolyticus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrrio.  
OX NCBI\_TaxID=670;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RMD 2210633 / Serotype O3:K6;  
RX MEDLINE=22508454; PubMed=12620739;  
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
RA Iijima Y., Nishima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
RA Yaenuga T., Honda T., Shingawa H., Hattori M., Iida T.;  
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism  
RT distinct from that of V. cholerae.";  
RL Lancet 361:743-749 (2003).  
DR EMBL; AP005088; BAC62537.1; --  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 287 AA; 34348 MW; 36799EAB97630E6 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 287;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMPE 8  
Db 226 FLOLMPE 233

RESULT 21  
Q7NGE4 PRELIMINARY; PRT; 295 AA.  
AC 07NGE4  
DT 01-MAR-2004 (TRENBLrel. 26, Created)  
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE G13227 protein.  
GN OrderedLocustNames=G13227;  
OS Gloeobacter violaceus.  
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.  
OX NCBI\_TaxID=33072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PCC 7421;  
RX MEDLINE=22977040; PubMed=14621292;  
RA Nakamura Y., Kaneo T., Sato S., Miumuro M., Miyashita H., Tsuchiya T.,  
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,  
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpō S.,

RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a  
RT cyanobacterium that lacks thylakoids."  
RL DNA Res. 10:137-145 (2003).  
CC -1- SIMILARITY: Contains 1 HTH arac/xylS-type DNA-binding domain.  
DR EMBL; AP006579; BAC91168.1; --  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0003700; F:transcription factor activity; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR000005; HTHARAC.  
DR Pfam; PF00165; HTH Arac; 2.  
DR PRINTS; PR00032; HTHARAC.  
DR PROSITE; PS00041; HTH\_ARAC\_FAMILY\_1; 2.  
DR PROSITE; PS01124; HTH\_ARAC\_FAMILY\_2; 1.  
KW Complete proteome; DNA-binding; Transcription regulation.  
SQ SEQUENCE 295 AA; 33473 MW; 77E3D14A5FB3B337 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 295;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMPEV 9  
Db 103 FLOLMPEV 111

RESULT 22  
Q6GB96 PRELIMINARY; PRT; 318 AA.  
AC 06GB96  
DT 05-JUL-2004 (TRENBLrel. 27, Created)  
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)  
DE FecCD transport family protein.  
GN ORFNames=SA50659;  
OS Staphylococcus aureus subsp. aureus MSSA476.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=282459;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MSSA476;  
RA Holden M.T.G., Fell B.J., Lindsay J.A., Peacock S.J., Day N.P.D.,  
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,  
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,  
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,  
RA Feltham D., Hance Z., Harris B., Hauser H., Holtroyd S., Jagsels K.,  
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,  
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K., Sanders M.,  
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrett B.G.,  
RA Spratt B.G., Parkhill J.;  
RT "Complete genomes of two clinical Staphylococcus aureus strains:  
RT evidence for the rapid evolution of virulence and drug resistance.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791 (2004).  
DR EMBL; BX571857; CAG42475.1; --  
DR InterPro; IPR000522; FecD.  
DR Pfam; PF01032; FecCD; 1.  
SQ SEQUENCE 318 AA; 35789 MW; 1B931B882739E3B CRC64;

Query Match 75.0%; Score 33; DB 2; Length 318;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLMPE 8  
Db 151 FLOLMPE 158

RESULT 23  
Q6G107 PRELIMINARY; PRT; 318 AA.  
AC 06G107  
DT 05-JUL-2004 (TRENBLrel. 27, Created)  
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)

```
DT 05-JUN-2004 (TREMBlrel. 27, last annotation update)
DE FeccD transport family protein.
GN Name=stb; ORFNames=SA0788;
OS Staphylococcus aureus subsp. aureus MRSA252.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=282458;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MRS252;
RA Holden M.T.G., Fell E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barton A.,
RA Bacon N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holtroyd S., Jagels K.,
RA James K.D., Leonard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL: BX571856; CAG39798.1; -
DR InterPro: IPR000522; FeccD.
DR Pfam: PF01032; FeccD; 1.
SQ SEQUENCE 318 AA; 35831 MW; C1B5A3BF19728BD6 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 318;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMEP 8
Db 151 FLOLIMP 158

RESULT 24
Q8NKM9 PRELIMINARY; PRT; 318 AA.
ID Q8NKM9
AC Q8NKM9
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE MM0696 protein.
GN OrderedLocustNames=MM0696;
OS Staphylococcus aureus (strain MM2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MM2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL: AP004824; BAB94561.1; -
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005215; P:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000522; FeccD.
DR Pfam: PF01032; FeccD; 1.
KM Complete proteome.
SQ SEQUENCE 318 AA; 35789 MW; 1B931B882735BE3B CRC64;

Query Match 75.0%; Score 33; DB 2; Length 318;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMEP 8
Db 151 FLOLIMP 158
```

```
RESULT 25
Q99VP0 PRELIMINARY; PRT; 318 AA.
ID Q99VP0
AC Q99VP0
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to ferrichrome ABC transporter permease.
GN OrderedLocustNames=SA0734;
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=158878;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hiramatsu H., Kubara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003360; BAB56896.1; -
DR FTR: G89845; G89845.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005215; P:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000522; FeccD.
DR Pfam: PF01032; FeccD; 1.
KM Complete proteome.
SQ SEQUENCE 318 AA; 35777 MW; 21ACE8A20E39BE21 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 318;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMEP 8
Db 151 FLOLIMP 158

RESULT 26
Q7A6S9 PRELIMINARY; PRT; 318 AA.
ID Q7A6S9
AC Q7A6S9
DT 05-JUN-2004 (TREMBlrel. 27, Created)
DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUN-2004 (TREMBlrel. 27, Last annotation update)
DE SA0689 protein.
GN OrderedLocustNames=SA0689;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=158879;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21311952; PubMed=11418146;
RX Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hiramatsu H., Kubara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003361; BAB41922.1; -
```

```

DR InterPro; IPR000522; FeCD.
RN Pfam; PF01032; FeCD; 1.
KW Complete proteome.
SQ SEQUENCE 318 AA; 35777 MW; 21ACE8A20E39BE21 CRC64;

Query Match
Best Local Similarity 75.0%; Score 33; DB 2; Length 318;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FLOLMMP 8
Db 151 FLOLMMP 158

RESULT 27
O8CPZ9 PRELIMINARY; PRT; 319 AA.
AC O8CPZ9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ferrichrome ABC transporter permease.
GN OrderedLocusNames=SE0516;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12560922;
RA Zhang Y.-Q., Ken S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
  Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
  Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
  "Genome-based analysis of virulence genes in a non-biofilm-forming
  RT Staphylococcus epidermidis strain (ATCC 12228).";
  RL Mol. Microbiol. 49:1577-1593(2003).
DR EMBL; AE016745; AA004113.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000522; FeCD.
DR Pfam; PF01032; FeCD; 1.
KW Complete proteome.
SQ SEQUENCE 319 AA; 35889 MW; 2C4BD99D1CA4F88 CRC64;

Query Match
Best Local Similarity 75.0%; Score 33; DB 2; Length 319;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLOLMMP 8
Db 152 FLOLMMP 159

RESULT 28
O7Z5K8 PRELIMINARY; PRT; 337 AA.
AC O7Z5K8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RAD23-like protein B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  CC Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX PubMed=15064313;
RA Huang X., Wang H., Xu M., Lu L., Xu Z., Li J., Zhou Z., Sha J.;
  "Expression of a Novel RAD23b mRNA Splice Variant in the Human
  RT Testis.";
```

```

RL J. Androl. 25:363-368(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Sha J.H., Zhou Z.M., Li J.M.;
  Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY313777; AAP81008.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STI1.
DR InterPro; IPR000449; UBA.
DR Pfam; PF00627; UBA; 2.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50030; UBA; 2.
SQ SEQUENCE 337 AA; 35033 MW; E9338A403D821F88 CRC64;
```

```

Query Match
Best Local Similarity 75.0%; Score 33; DB 2; Length 337;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```

Oy 1 FLOLMMPV 9
Db 252 FLOLMMPV 260
```

```

RESULT 29
O84LJ2 PRELIMINARY; PRT; 366 AA.
AC O84LJ2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RAD23-like protein.
GN Name=AtRAD23-2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
  OX eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=flower bud;
RA Ishikawa Y., Endo M., Abe K., Osakabe K., Nakajima N., Saji H.,
  Ito Y., Ichikawa H., Kameya T., Toki S.;
  Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB109195; BAC76391.1; -.
DR HSSP; P54725; IIFY.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR InterPro; IPR009020; Prot_inh_propept.
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STI1.
DR InterPro; IPR000449; UBA.
DR InterPro; IPR009060; UBA_like.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; Ubiquitin; 1.
DR PRINTS; PR01839; RAD23PROTEIN.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PS50030; UBA; 2.
DR PROSITE; PS50053; UBQUITIN_2; 1.
SQ SEQUENCE 366 AA; 39626 MW; 5ECA2D1604B2B001 CRC64;
```

```

Query Match
Best Local Similarity 75.0%; Score 33; DB 2; Length 366;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```



Qy 1 FLOLMEPV 8  
Db 287 FLOLMEPV 294

RESULT 30  
Q6TLDO PRELIMINARY; PRT; 380 AA.  
AC Q6TLDO; 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DE RAD23 homolog B.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
NCBI\_TaxID=7955;  
[1]  
SEQUENCE FROM N.A.  
RC TISSUE=Kidney marrow;  
RA Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,  
RA Sheng Y., Chen Y., Ruan Z., Jiang C.L., Fan H.Y., Zou L.I.,  
RA Kanki J.P., Look A.T., Chen Z.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY394976; AAQ94603.1;  
DR InterPro; IPR004806; Rad23.  
DR InterPro; IPR006366; STI1.  
DR InterPro; IPR009060; UBA\_like.  
DR InterPro; IPR006261; Ubiqutin.  
DR Pfam; PF00240; ubi\_qutin; 1.  
DR PRINTS; PR01839; RAD23PROTEIN.  
DR SMART; SM00727; STI1; 1.  
DR SMART; SM00165; UBA; 2.  
DR SMART; SM00213; UBQ; 1.  
DR TIGRFAMs; TIGR00601; rad23; 1.  
DR PROSITE; PSS0030; UBA; 2.  
DR PROSITE; PSS0053; UBIQUITIN\_2; 1.  
SQ SEQUENCE 380 AA; 40107 MW; 5D05919718B28652 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 380;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 300 FLOLMEPV 308

RESULT 31  
AAQ94603 PRELIMINARY; PRT; 380 AA.  
AC AAQ94603;  
DT 02-MAR-2004 (TREMBLrel. 27, Created)  
DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)  
DE RAD23 homolog B.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
NCBI\_TaxID=7955;  
[1]  
SEQUENCE FROM N.A.  
RC TISSUE=Kidney marrow;  
RA Song H.D., Wu X.Y., Sun X.J., Zhou Y., Liu T.X., Deng M., Zhang G.W.,  
RA Sheng Y., Chen Y., Ruan Z., Jiang C.L., Fan H.Y., Zou L.I.,  
RA Kanki J.P., Look A.T., Chen Z.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY394976; AAQ94603.1; -.

SO SEQUENCE 380 AA; 40107 MW; 5D05919718B28652 CRC64;  
Query Match 75.0%; Score 33; DB 2; Length 380;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOLMEPV 9  
Db 300 FLOLMEPV 308

RESULT 32  
Q6PHE9 PRELIMINARY; PRT; 382 AA.  
AC Q6PHE9;  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DE RAD23 homolog B.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
NCBI\_TaxID=7955;  
[1]  
SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RC MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalek U., Smailus D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC056578; AAH56578.1; -.  
DR InterPro; IPR004806; Rad23.  
DR InterPro; IPR006366; STI1.  
DR InterPro; IPR000448; UBA.  
DR InterPro; IPR009060; UBA\_like.  
DR InterPro; IPR006261; Ubiqutin.  
DR Pfam; PF00240; ubi\_qutin; 1.  
DR Pfam; PF00627; UBA; 2.  
DR PRINTS; PR01839; RAD23PROTEIN.  
DR SMART; SM00727; STI1; 1.  
DR SMART; SM00165; UBA; 2.  
DR SMART; SM00213; UBQ; 1.  
DR TIGRFAMs; TIGR00601; rad23; 1.  
DR PROSITE; PSS0030; UBA; 2.  
DR PROSITE; PSS0053; UBIQUITIN\_2; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 382 AA; 40308 MW; 7BCD165BEBBC8 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 382;

Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLOLMPEPV 9  
Db 302 FLOLMPEPV 310

RESULT 33  
AAH56578 PRELIMINARY; PRT; 382 AA.  
AC AAH56578;  
DT 24-MAY-2004 (TREMBLrel. 27, Created)  
DT 24-MAY-2004 (TREMBLrel. 27, Last sequence update)  
DE Hypothetical protein zgc:65951.  
GN ZGC:65951.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
FAhey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
Krzyniński M.I., Skalska U., Smalits D.E., Schnerch A., Schein J.B.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC056578; AAH56578.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 382 AA; 40308 MW; 7BEC3D165BEBEC8 CRC64;  
QY Query Match 75.0%; Score 33; DB 2; Length 382;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Db 1 FLOLMPEPV 9  
302 FLOLMPEPV 310  
RESULT 34  
R23B\_HUMAN STANDARD; PRT; 409 AA.  
ID R23B\_HUMAN  
AC P54727; Q8WUB0;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-2004 (Rel. 45, Last annotation update)  
DE UV excision repair protein RAD23 homolog B (hHR23B) (XP-C repair  
complementing complex 58 kDa protein) (p58).

GN Name=RAD23B;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=9422030; PubMed=8168482;  
RA Masutani C., Sugawara K., Yanagisawa J., Sonoyama T., Uli M.,  
RA Enomoto T., Takio K., Tanaka K., van der Spek P.J., Bootama D.,  
RA Hoeljmakers J.H.J., Hanaoka F.;  
RT "Purification and cloning of a nucleotide excision repair complex  
involving the xeroderma pigmentosum group C protein and a human  
homologue of yeast RAD23.";  
RL EMBO J. 13:1831-1843(1994).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANT VAL-249.  
RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,  
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,  
RA Schackwitz W.S., Sherwood J.K., Wiltrik L.A., Nickerson D.A.;  
RT "NIHS-SNPs, environmental genome project, NIHS ES15478, Department  
of Genome Sciences, Seattle, WA (URL: <http://egp.gs.washington.edu>).";  
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Ramsey H.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A., AND VARIANT VAL-249.  
RC TISSUE=Testis;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
FAhey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzyniński M.I., Skalska U., Smalits D.E.,  
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [5]  
RP FUNCTION: Involved in DNA excision repair. May play a part in DNA  
damage recognition and/or in altering chromatin structure to allow  
access by damage-processing enzymes.  
CC -1- SUBUNIT: Heterodimer of a 125 kDa subunit (p125) and of a 58 kDa  
subunit (p58). Interacts with MJD.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
CC -1- DOMAIN: The ubiquitin-like domain mediates interaction with MJD.  
CC -1- SIMILARITY: Contains 2 UBA domains.  
CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
DR EMBL; D21090; BAA04652.1; -.  
DR EMBL; A1137852; CAD1375.1; -.  
DR EMBL; A1165178; AAN47194.1; -.  
DR EMBL; BC020973; AAN20973.1; -.

DR PIR; S44346; S44346.  
 DR PDB; IUEI; NMR; A=1-95.  
 DR OGP; P54727; -.  
 DR Genew; HGNC:9813; RAD23B.  
 DR Reactome; P54727; -.  
 DR MIM; 600062; -.  
 DR GO; GO:0005634; C:nucleus; TAS.  
 DR GO; GO:0003697; F:single-stranded DNA binding; TAS.  
 DR GO; GO:0006289; P:nucleotide-excision repair; TAS.  
 DR InterPro; IPR004806; Rad23.  
 DR InterPro; IPR006635; STI1.  
 DR InterPro; IPR000449; UBA.  
 DR InterPro; IPR009060; UBA\_like.  
 DR Pfam; PF00627; UBA; 2.  
 DR Pfam; PF00240; Ubiquitin; 1.  
 DR PRINTS; PR01839; RAD23PROTEIN.  
 DR SMART; SM00727; STI1; 1.  
 DR SMART; SM00165; UBA; 2.  
 DR SMART; SM00213; UBO; 1.  
 DR TIGRFAMs; TIGR00601; rad23; 1.  
 DR PROSITE; PSS0030; UBA; 2.  
 DR PROSITE; PSS0053; UBIQUITIN\_2; 1.  
 DR 3D-structure; Direct protein sequencing; DNA damage; DNA repair;  
 KW Nuclear protein; Polymorphism; Repeat.  
 FT DOMAIN 1 79 Ubiquitin-like.  
 FT DOMAIN 188 228 UBA 1.  
 FT DOMAIN 364 404 UBA 2.  
 FT DOMAIN 103 106 Poly-Thr.  
 FT DOMAIN 254 260 Poly-Ala.  
 FT DOMAIN 261 269 Poly-Thr.  
 FT DOMAIN 336 348 Poly-Gly.  
 FT VARIANT 249 249 A -> V (in dbSNP:18053329).  
 FT VARIANT 249 249 /FTID=VAR\_014350.  
 SQ SEQUENCE 409 AA; 43171 MW; C026C78273BCB289 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 409;  
 Best Local Similarity 66.7%; Pred. No. 2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOQLMEPV 9  
 Db 324 FLOQLMEPV 332

RESULT 35  
 ID Q6NVC3 PRELIMINARY; PRT; 415 AA.  
 AC Q6NVC3;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Rad23b protein.  
 GN Name=Rad23b;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Tothiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Viallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzyzanski M.I., Skalek U., Smalms D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Maitra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC068193; AA68193.1; -.  
 DR GO; GO:0006974; P:response to DNA damage stimulus; IDA.  
 DR GO; GO:0007283; P:spermatogenesis; IMP.  
 DR InterPro; IPR004806; Rad23.  
 DR InterPro; IPR006635; STI1.  
 DR InterPro; IPR000449; UBA.  
 DR InterPro; IPR009060; UBA\_like.  
 DR Pfam; PF00627; UBA; 2.  
 DR Pfam; PF00240; Ubiquitin; 1.  
 DR PRINTS; PR01839; RAD23PROTEIN.  
 DR SMART; SM00727; STI1; 1.  
 DR SMART; SM00165; UBA; 2.  
 DR SMART; SM00213; UBO; 1.  
 DR TIGRFAMs; TIGR00601; rad23; 1.  
 DR PROSITE; PSS0030; UBA; 2.  
 DR PROSITE; PSS0053; UBIQUITIN\_2; 1.  
 SQ SEQUENCE 415 AA; 43415 MW; 3C53336BC784786B CRC64;

Query Match 75.0%; Score 33; DB 2; Length 415;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOQLMEPV 9  
 Db 323 FLOQLMEPV 331

RESULT 36  
 ID AAH68193 PRELIMINARY; PRT; 415 AA.  
 AC AAH68193;  
 DT 14-APR-2004 (TrEMBLrel. 27, Created)  
 DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 14-APR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Rad23b protein.  
 GN Name=Rad23b;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Tothiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Viallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywniuk M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Maria M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strauberg R.;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC068193; AAH68193.1; -  
 SQ SEQUENCE 415 AA; 43415 MW; 3C53336BC784786B CRC64;

Query Match 75.0%; Score 33; DB 2; Length 415;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
 DB 323 FIOMLNEPV 331

RESULT 37  
 R23B\_MOUSE STANDARD; PRT; 416 AA.  
 ID R23B\_MOUSE  
 AC P54728;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE UV excision repair protein RAD23 homolog B (hHR23B) (Xp-C repair  
 DE complementing complex 58 kDa protein) (p58).  
 GN Name=Rad23b; Synonyms=Mhr23b;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Testis;  
 RX MEDLINE=96403997; PubMed=8808275;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uscid T.B., Toshiyuki S., Carninci P., Priege C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muliyil S.J.,  
 RA Bork S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywniuk M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Maria M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Involved in DNA excision repair. May play a part in DNA  
 CC damage recognition and/or in altering chromatin structure to allow

CC access by damage-processing enzymes.  
 CC -1- SUBUNIT: Heterodimer of a 125 kDa subunit (p125) and of a 58 kDa  
 CC subunit (p58).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: Contains 2 UBA domains.  
 CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X92411; CAA63146.1; -  
 CC EMBL: BC027747; AAH27747.1; -  
 CC HSSP: P54725; IDV0.  
 CC MGD: MGI:105128; Rad23b.  
 CC GO: GO:007283; P.spermatogenesis; IMP.  
 CC InterPro: IPR004806; Rad23.  
 CC InterPro: IPR006636; Still.  
 CC InterPro: IPR000449; UBA.  
 CC InterPro: IPR009060; UBA-like.  
 CC InterPro: IPR000626; Ubiquitin.  
 CC Pfam: PF00627; UBA; 2.  
 CC Pfam: PF00240; ubiquitin; 1.  
 CC PRINTS: PR01839; RAD23PROTEIN.  
 CC SMART: SM00727; STIL; 1.  
 CC SMART: SM00165; UBA; 2.  
 CC SMART: SM00213; UBC; 1.  
 CC TIGRfam: TIGR00601; rad23; 1.  
 CC PROSITE: PS50053; UBA; 2.  
 CC PROSITE: PS50053; Ubiquitin\_2; 1.  
 CC DNA damage; DNA repair; Nuclear protein; Repeat.  
 FT DOMAIN 1 79 Ubiquitin-like.  
 FT DOMAIN 2 228 UBA 1.  
 FT DOMAIN 3 371 UBA 2.  
 FT DOMAIN 4 411 UBA 2.  
 FT DOMAIN 5 255 Poly-Ala.  
 FT DOMAIN 6 262 Poly-Thr.  
 FT DOMAIN 7 355 Poly-Gly.  
 SQ SEQUENCE 416 AA; 43516 MW; 1380245A6D892205 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 416;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
 DB 324 FIOMLNEPV 332

RESULT 38  
 O6FUT0 PRELIMINARY; PRT; 435 AA.  
 ID O6FUT0  
 AC O6FUT0;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Similarities with tr|Q03983 Saccharomyces cerevisiae YDR179wa.  
 GN ORFNames=CGL0R058195;  
 OS Candida glabrata (Yeast) (Torulopsis glabrata).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5478;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS138;  
 RG GENOLEVRES;  
 RA Dujon B., Sherman D., Fischer G., Durrans P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marcq C., Neuveglise C., Talla E.,  
 RA Goffard N., Frangeul L., Aigle M., Anhouard V., Babour A., Barbe V.,  
 RA Bernay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,

RA Boisrame A., Boyer J., Cactolico L., Confanioleri F., de Darvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Gropi A.,  
 RA Hantoune F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicand J.M., Nikolaki M., Oztas S., Ozler-Kalogeropoulos O.,  
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Sweeney D., Tekala F., Wesolowski-Jouvel M., Westhof E., Wirth B.,  
 RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,  
 RA Wincker P., Soucier J.L.,  
 RT "genome evolution in yeasts.",  
 RL Nature 430:35-44(2004).  
 DR ENBL: CR380952; CAG59148.1; -.  
 SQ SEQUENCE 435 AA; 50090 MW; 404285E838181F47 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 435;  
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
 DB 418 FLOLMPEPI 426

RESULT 39  
 ID 08S012 PRELIMINARY; PRT; 1569 AA.  
 AC 08S012;  
 DT 01-JUN-2002 (TEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)  
 DE Putative glucan synthase.  
 GN Name=OJ1619 P12.14;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Oryzoideae; Oryzaceae; Oryza.  
 NC NCB1\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Maekawa M., Arikawa K., Chiden Y., Hayashi M.,  
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
 RA Hijioka S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
 RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuuchi A., Kamiya K.,  
 RA Karsawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
 RA Machida K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Sait S., Sakai K., Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
 RA Zhong H., Yamaoka H., Endo T., Ito H., Hain J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.,  
 RT "The genome sequence and structure of rice chromosome 1.",  
 RL Nature 420:312-316(2002).  
 DR ENBL: AP003447; BAB90326.1; -.  
 DR Gramene: O8S012;  
 DR GO: GO:0000148; C:1,3-beta-glucan synthase complex; IEA.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0003843; F:1,3-beta-glucan synthase activity; IEA.  
 DR GO: GO:0006075; F:beta-1,3 glucan biosynthesis; IEA.  
 DR InterPro: IPR003440; Glyco trans 48.  
 DR Pfam: PF02364; Glucan synthase; 2.  
 DR SEQUENCE 1569 AA; 181963 MW; 436A2A8ED57408A1 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 1569;  
 Best Local Similarity 55.6%; Pred. No. 8.3e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLMPEPV 9  
 DB 416 FLOLMPEPI 424

RESULT 40  
 ID 08IMS2 PRELIMINARY; PRT; 2746 AA.  
 AC 08IMS2;  
 DT 01-MAR-2003 (TEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)  
 DE CG8318-PC.  
 GN Name=Nfi; ORFNames=CG8318;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCB1\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,  
 RA Adair J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evansgelist C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of Drosophila melanogaster.",  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426055; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
 melanogaster euchromatic genome sequence.",  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,

RA Patel S., Friese B., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomic perspective.";   
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu I., Berman B.P.,  
 RA Betencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";   
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FLYBASE;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FLYBASE;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003754; AAN14067.2; -.  
 DR HSSP; P21359; INF1.  
 DR FLYBase; FBgn0015269; NF1.  
 DR GO; GO:0016049; P:cell growth; NAS.  
 DR GO; GO:0007611; P:learning and/or memory; NAS.  
 DR GO; GO:0045475; P:locomotor rhythm; NAS.  
 DR GO; GO:0007622; P:rhythmic behavior; NAS.  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR001251; CRAL\_TRIO\_C.  
 DR InterPro; IPR001936; RasGAP.  
 DR Pfam; PF00616; RasGAP; 1.  
 DR SMART; SM00323; RasGAP; 1.  
 DR SMART; SM00516; SEC14; 1.  
 DR PROSITE; PS50191; CRAL\_TRIO; 1.  
 DR PROSITE; PS50018; RAS\_GTPASE\_ACTIV\_2; 1.  
 SO SEQUENCE 2746 AA; 311107 MW; 5D24FB68540A0CC8 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 2746;  
 Best Local Similarity 55.6%; Pred.No. 1.5e+03;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQILMEPV 9  
 :||:|:|:  
 Db 1340 YLQMLLEPL 1348

Search completed: January 12, 2005, 20:14:02  
 Job time : 93 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:46:44 ; Search time 21.6 Seconds  
(without alignments)  
27.633 Million cell updates/sec

Title: US-09-870-216C-5  
Perfect score: 44  
Sequence: 1 FLOLEFPAV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep: \*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep: \*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep: \*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep: \*  
5: /cgn2\_6/prodata/1/iaa/POCUS\_COMB.pep: \*  
6: /cgn2\_6/prodata/1/iaa/backfillseq.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	75.0	420	4	US-09-248-796A-18548
2	32	72.7	419	4	US-09-252-991A-27064
3	32	72.7	523	4	US-09-270-767-43156
4	32	72.7	633	4	US-09-919-060-13
5	32	72.7	2289	3	US-09-051-019-2
6	31	70.5	210	1	US-07-667-276A-2
7	31	70.5	287	4	US-09-543-681A-7259
8	31	70.5	379	4	US-09-270-767-32891
9	31	70.5	379	4	US-09-270-767-48108
10	31	70.5	524	4	US-09-252-991A-17710
11	31	70.5	803	4	US-09-489-039A-12742
12	30	68.2	191	4	US-09-621-976-6999
13	30	68.2	299	4	US-09-584-568C-6
14	30	68.2	300	4	US-09-194-146-6
15	30	68.2	333	4	US-09-248-796A-17292
16	30	68.2	370	3	US-09-134-001C-3403
17	30	68.2	370	4	US-09-710-279-696
18	30	68.2	370	4	US-09-710-279-1328
19	30	68.2	582	4	US-09-516-914-21
20	30	68.2	1176	4	US-09-976-594-793
21	29	65.9	57	4	US-09-513-999C-7337
22	29	65.9	130	3	US-09-102-769-29
23	29	65.9	134	4	US-09-270-767-60072
24	29	65.9	141	4	US-09-621-976-6981
25	29	65.9	141	4	US-09-621-976-6982
26	29	65.9	141	4	US-09-513-999C-7895
27	29	65.9	178	4	US-09-153-447-21

28	29	65.9	185	4	US-09-489-039A-9443	Sequence 9443, Ap
29	29	65.9	212	4	US-09-328-352-7485	Sequence 7485, Ap
30	29	65.9	242	2	US-08-622-352A-3	Sequence 3, Appl
31	29	65.9	242	3	US-08-826-390-3	Sequence 3, Appl
32	29	65.9	316	4	US-09-540-236-3467	Sequence 3467, Ap
33	29	65.9	389	4	US-09-270-767-44624	Sequence 44624, A
34	29	65.9	427	4	US-09-252-991A-19612	Sequence 19612, A
35	29	65.9	467	4	US-09-107-532A-6186	Sequence 6186, Ap
36	29	65.9	512	4	US-09-107-532A-7248	Sequence 7248, Ap
37	29	65.9	642	4	US-09-252-991A-19386	Sequence 19386, A
38	29	65.9	680	2	US-08-674-351-2	Sequence 2, Appl
39	29	65.9	924	4	US-09-267-311-2	Sequence 2, Appl
40	29	65.9	1024	3	US-09-091-117-5	Sequence 5, Appl
41	29	65.9	1338	4	US-09-248-796A-16051	Sequence 16051, A
42	29	65.9	1596	4	US-09-538-092-887	Sequence 887, Ap
43	28	63.6	51	3	US-09-177-249-289	Sequence 289, Ap
44	28	63.6	70	4	US-09-248-796A-27334	Sequence 27334, A
45	28	63.6	72	4	US-09-663-600A-93	Sequence 93, Appl

#### ALIGNMENTS

```

RESULT 1
US-09-248-796A-18548
; Sequence 18548, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18548
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18548

Query Match      75.0%; Score 33; DB 4; Length 420;
Best Local Similarity 85.7%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      3 QLEFPAV 9
DB      339 QLEFPAI 345

RESULT 2
US-09-252-991A-27064
; Sequence 27064, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27064
; LENGTH: 419
; TYPE: PRT

```

ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27064

Query Match  
Best Local Similarity 72.7%; Score 32; DB 4; Length 419;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLOLEFPAV 9  
DB 226 FLOLEFPAV 233

RESULT 3  
US-09-270-767-43156  
; Sequence 43156, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 43156  
; LENGTH: 523  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
US-09-270-767-43156  
OTHER INFORMATION: Xaa means any amino acid

Query Match  
Best Local Similarity 72.7%; Score 32; DB 4; Length 523;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFPA 8  
DB 477 FLOLEFPA 484

RESULT 4  
US-09-919-060-13  
; Sequence 13, Application US/09919060  
; Patent No. 6638744  
; GENERAL INFORMATION:  
; APPLICANT: Wisniewski, Nancy  
; TITLE OF INVENTION: CANINE COX-1 AND COX-2 NUCLEIC ACID MOLECULES, PROTEINS AND USES  
; FILE REFERENCE: AD-1  
; CURRENT APPLICATION NUMBER: US/09/919,060  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: 60/224,486  
; PRIOR FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 13  
; LENGTH: 633  
; TYPE: PRT  
; ORGANISM: Canis familiaris  
US-09-919-060-13

Query Match  
Best Local Similarity 72.7%; Score 32; DB 4; Length 633;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFPA 7  
DB 389 FLOLEFPA 395

RESULT 5  
US-09-051-019-2

; Sequence 2, Application US/09051019  
; Patent No. 6103229  
; GENERAL INFORMATION:  
; APPLICANT: RAHMANN, Regine and QUADBECK-SEEGER, Claudia  
; TITLE OF INVENTION: Regulatory gene from *Ustilago maydis*  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Keil & Weinkeuf  
; STREET: 1101 Connecticut Avenue  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.2 Mb storage  
; COMPUTER: IBM AT-compatible, Pentium processor  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: WordPerfect version 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/051,019  
; FILING DATE: 31-MAR-1998  
; CLASSIFICATION:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2289 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-09-051-019-2

Query Match  
Best Local Similarity 72.7%; Score 32; DB 3; Length 2289;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEFPAV 9  
DB 1510 FLOLEFPAV 1518

RESULT 6  
US-07-667-276A-2  
; Sequence 2, Application US/07667276A  
; Patent No. 5470971  
; GENERAL INFORMATION:  
; APPLICANT: Kondo, Keiji  
; TITLE OF INVENTION: Inouye, Masayori  
; TITLE OF INVENTION: STRESS-INDUCED PROTEINS, GENES CODING  
; TITLE OF INVENTION: THEREFOR, TRANSFORMED CELLS OF ORGANISMS, METHODS AND  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weiser & Associates  
; STREET: 230 S. Fifteenth Street, Suite 500  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/667,276A  
; FILING DATE: 11-MAR-1991  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weiser, Gerard J.  
; REGISTRATION NUMBER: 19,763  
; REFERENCE/DOCKET NUMBER: 377,5351P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-875-8383  
; TELEFAX: 215-875-8394  
; INFORMATION FOR SEQ ID NO: 2:



SEQUENCE CHARACTERISTICS:  
LENGTH: 210 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-667-276A-2

Query Match 70.5%; Score 31; DB 1; Length 210;  
Best Local Similarity 55.6%; Pred. No. 97;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
|:|:|:|:  
Db 78 FSELDPDAI 86

RESULT 7  
US-09-543-681A-7259  
Sequence 7259, Application US/09543681A  
Patent No. 6605709

GENERAL INFORMATION:  
APPLICANT: GARY BRETON  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
FILE REFERENCE: 2709.1002-001  
CURRENT APPLICATION NUMBER: US/09/543,681A  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 60/128,706  
PRIOR FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 8344  
SEQ ID NO 7259  
LENGTH: 287  
TYPE: PRT  
ORGANISM: Proteus mirabilis  
US-09-543-681A-7259

Query Match 70.5%; Score 31; DB 4; Length 287;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QLEPDAV 9  
|:|:|:|:  
Db 139 QVEPDAV 145

RESULT 8  
US-09-270-767-32891  
Sequence 32891, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 32891  
LENGTH: 379  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-270-767-32891

Query Match 70.5%; Score 31; DB 4; Length 379;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
|:|:|:|:  
Db 360 FLSMOYDAV 368

RESULT 9

US-09-270-767-48108  
Sequence 48108, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 48108  
LENGTH: 379  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-270-767-48108

Query Match 70.5%; Score 31; DB 4; Length 379;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
|:|:|:|:  
Db 360 FLSMOYDAV 368

RESULT 10  
US-09-252-991A-17710  
Sequence 17710, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 17710  
LENGTH: 524  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17710

Query Match 70.5%; Score 31; DB 4; Length 524;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
|:|:|:|:  
Db 245 FLOVEYDAL 253

RESULT 11  
US-09-489-039A-12742  
Sequence 12742, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 12742  
LENGTH: 803

TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12742

Query Match  
Best Local Similarity 70.5%; Score 31; DB 4; Length 803;  
Best Local Similarity 77.8%; Pred. No. 3.8e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
DB 577 FLOLEPDAV 585

RESULT 12  
US-09-621-976-6999  
Sequence 6999, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 6999  
LENGTH: 191  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-621-976-6999

Query Match  
Best Local Similarity 68.2%; Score 30; DB 4; Length 191;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FLOLEPDAV 7  
DB 25 FLOLEPDAV 30

RESULT 13  
US-09-584-568C-6  
Sequence 6, Application US/09584568C  
Patent No. 6500657  
GENERAL INFORMATION:  
APPLICANT: Glucksmann, Maria, Alexandra et al.  
TITLE OF INVENTION: 33167, A NOVEL HUMAN HYDROLASE AND USES THEREFOR  
FILE REFERENCE: NMI-140  
CURRENT APPLICATION NUMBER: US/09/584,568C  
CURRENT FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: 60/193,954  
PRIOR FILING DATE: 2000-03-31  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 299  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-09-584-568C-6

Query Match  
Best Local Similarity 68.2%; Score 30; DB 4; Length 299;  
Best Local Similarity 55.6%; Pred. No. 2.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
DB 266 FLOLEPDAV 274

RESULT 14  
US-09-194-146-6

Sequence 6, Application US/09194146  
Patent No. 6458557  
GENERAL INFORMATION:  
APPLICANT: Miller, Brian  
APPLICANT: Diaz-Torres, Maria  
TITLE OF INVENTION: Gram-Positive Microorganism Formate Pathway  
FILE REFERENCE: GC395-US  
CURRENT APPLICATION NUMBER: US/09/194,146  
CURRENT FILING DATE: 1998-11-19  
PRIOR APPLICATION NUMBER: GB 9724627.6  
PRIOR FILING DATE: 1997-11-20  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 6  
LENGTH: 300  
TYPE: PRT  
ORGANISM: Bacillus  
US-09-194-146-6

Query Match  
Best Local Similarity 68.2%; Score 30; DB 4; Length 300;  
Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 7  
DB 63 FLOLEPDAV 69

RESULT 15  
US-09-248-796A-17292  
Sequence 17292, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 17292  
LENGTH: 333  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-17292

Query Match  
Best Local Similarity 68.2%; Score 30; DB 4; Length 333;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 6  
DB 129 FLOLEPDAV 134

RESULT 16  
US-09-134-001C-3403  
Sequence 3403, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 3403  
 ; LENGTH: 370  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis  
 US-09-134-001C-3403

Query Match 68.2%; Score 30; DB 3; Length 370;  
 Best Local Similarity 75.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEPDV 9  
 DB 5 LQAQPDV 12

RESULT 17  
 US-09-710-279-696  
 ; Sequence 696, Application US/09710279  
 ; Patent No. 6703492  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIMBERLY, WILLIAM JOHN  
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
 ; FILE REFERENCE: PU3480US  
 ; CURRENT FILING DATE: 2000-11-09  
 ; PRIOR APPLICATION NUMBER: 60/164,258  
 ; PRIOR FILING DATE: 1999-11-09  
 ; NUMBER OF SEQ ID NOS: 4472  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 696  
 ; LENGTH: 370  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
 US-09-710-279-696

Query Match 68.2%; Score 30; DB 4; Length 370;  
 Best Local Similarity 75.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEPDV 9  
 DB 5 LQAQPDV 12

RESULT 18  
 US-09-710-279-1328  
 ; Sequence 1328, Application US/09710279  
 ; Patent No. 6703492  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIMBERLY, WILLIAM JOHN  
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
 ; FILE REFERENCE: PU3480US  
 ; CURRENT APPLICATION NUMBER: US/09/710,279  
 ; CURRENT FILING DATE: 2000-11-09  
 ; PRIOR APPLICATION NUMBER: 60/164,258  
 ; PRIOR FILING DATE: 1999-11-09  
 ; NUMBER OF SEQ ID NOS: 4472  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 1328  
 ; LENGTH: 370  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
 US-09-710-279-1328

Query Match 68.2%; Score 30; DB 4; Length 370;

Best Local Similarity 75.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEPDV 9  
 DB 5 LQAQPDV 12

RESULT 19  
 US-09-516-914-21  
 ; Sequence 21, Application US/09516914  
 ; Patent No. 6333401  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Breinig, Sabine  
 ; APPLICANT: Fuchs, Georg  
 ; TITLE OF INVENTION: Phenol-induced Proteins of Thauera aromatica  
 ; FILE REFERENCE: BC1006 US NA  
 ; CURRENT APPLICATION NUMBER: US/09/516,914  
 ; CURRENT FILING DATE: 2000-03-01  
 ; EARLIER APPLICATION NUMBER: 60/122,952  
 ; EARLIER FILING DATE: 1999-03-05  
 ; NUMBER OF SEQ ID NOS: 44  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 21  
 ; LENGTH: 582  
 ; TYPE: PRT  
 ; ORGANISM: Thauera aromatica  
 US-09-516-914-21

Query Match 68.2%; Score 30; DB 3; Length 582;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLEPD 7  
 DB 109 LQLEPD 114

RESULT 20  
 US-09-976-594-793  
 ; Sequence 793, Application US/09976594  
 ; Patent No. 6673549  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Furness, Michael  
 ; APPLICANT: Buchbinder, Jenny  
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
 ; FILE REFERENCE: PA-0041 US  
 ; CURRENT APPLICATION NUMBER: US/09/976,594  
 ; CURRENT FILING DATE: 2001-10-12  
 ; PRIOR APPLICATION NUMBER: 60/240,409  
 ; PRIOR FILING DATE: 2000-10-12  
 ; NUMBER OF SEQ ID NOS: 1143  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 793  
 ; LENGTH: 1176  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; NAME/KEY: misc.feature  
 ; OTHER INFORMATION: Incyte ID No. 6673549 1806212CD1  
 US-09-976-594-793

Query Match 68.2%; Score 30; DB 4; Length 1176;  
 Best Local Similarity 100.0%; Pred. No. 8.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLEPD 7  
 DB 1010 LQLEPD 1015

RESULT 21  
 US-09-513-999C-7337

; Sequence 7337, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7337  
; LENGTH: 57  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-513-999C-7337

Query Match 65.9%; Score 29; DB 4; Length 57;  
Best Local Similarity 75.0%; Pred. No. 64;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLEFPA 8  
|||  
Db 32 FLELEMDA 39

RESULT 22  
US-09-302-769-29  
; Sequence 29, Application US/09302769  
; Patent No. 6323317  
; GENERAL INFORMATION:  
; APPLICANT: HILTON, Douglas J  
; APPLICANT: ALEXANDER, Warren S  
; APPLICANT: VINEY, Elizabeth M  
; APPLICANT: WILSON, Tracey A  
; APPLICANT: RICHARDSON, Rachael T  
; APPLICANT: STARR, Robyn  
; APPLICANT: NICHOLSON, Sandra E  
; APPLICANT: METCALF, Donald  
; APPLICANT: NICOLA, Nicos A  
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS  
; FILE REFERENCE: 109762  
; CURRENT APPLICATION NUMBER: US/09/302,769  
; CURRENT FILING DATE: 1999-04-30  
; PRIOR APPLICATION NUMBER: 08/962,560  
; PRIOR FILING DATE: 1997-10-31  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 29  
; LENGTH: 130  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-302-769-29

Query Match 65.9%; Score 29; DB 3; Length 130;  
Best Local Similarity 71.4%; Pred. No. 1.5e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFPA 7  
|||  
Db 120 FLOLEFPA 126

RESULT 23  
US-09-270-767-60072  
; Sequence 60072, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 60072  
; LENGTH: 134  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-60072

Query Match 65.9%; Score 29; DB 4; Length 134;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEFPAV 9  
|||  
Db 87 LQLEFPAV 94

RESULT 24  
US-09-621-976-6981  
; Sequence 6981, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6981  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-621-976-6981

Query Match 65.9%; Score 29; DB 4; Length 141;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEFPAV 9  
|||  
Db 79 LQLEFPAV 86

RESULT 25  
US-09-621-976-6982  
; Sequence 6982, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6982  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-621-976-6982

Query Match 65.9%; Score 29; DB 4; Length 141;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEPDAV 9  
|:|:|:|  
Db 79 LQMEFRAV 86

RESULT 26  
US-09-513-999C-7895  
; Sequence 7895, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59, US2, REG  
; CURRENT APPLICATION NUMBER: US/09/513, 999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO: 7895  
; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-513-999C-7895

Query Match 65.9%; Score 29; DB 4; Length 141;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEPDAV 9  
|:|:|:|  
Db 79 LQMEFRAV 86

RESULT 27  
US-09-153-447-21  
; Sequence 21, Application US/09153447  
; Patent No. 6610838  
; GENERAL INFORMATION:  
; APPLICANT: BERGSTROM, Sven  
; TITLE OF INVENTION: P13 ANTIGENS FROM BORRELLIA  
; FILE REFERENCE: 454312-1130.1  
; CURRENT APPLICATION NUMBER: US/09/153,447  
; CURRENT FILING DATE: 1998-09-15  
; EARLIER APPLICATION NUMBER: 60/059,036  
; EARLIER FILING DATE: 1997-09-10  
; EARLIER APPLICATION NUMBER: 1041/97  
; EARLIER FILING DATE: 1997-09-10  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO: 21  
; LENGTH: 178  
; TYPE: PRT  
; ORGANISM: Borrelia afzelii  
US-09-153-447-21

Query Match 65.9%; Score 29; DB 4; Length 178;  
Best Local Similarity 77.8%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLQLEPDAV 9  
|:|:|:|  
Db 72 FLILGPDV 80

RESULT 28  
US-09-489-039A-9443  
; Sequence 9443, Application US/09489039A  
; Patent No. 6610836

; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709,2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO: 9443  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9443

Query Match 65.9%; Score 29; DB 4; Length 185;  
Best Local Similarity 62.5%; Pred. No. 2.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLQLEPDA 8  
|:|:|:|  
Db 170 FMQLQFDA 177

RESULT 29  
US-09-328-352-7485  
; Sequence 7485, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO: 7485  
; LENGTH: 212  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7485

Query Match 65.9%; Score 29; DB 4; Length 212;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLQLEPDAV 9  
|:|:|:|  
Db 88 FLQLSFPAL 96

RESULT 30  
US-08-622-352A-3  
; Sequence 3, Application US/08622352A  
; Patent No. 5824546  
; GENERAL INFORMATION:  
; APPLICANT: Bishai, William R.  
; APPLICANT: Demajo, James  
; TITLE OF INVENTION: REGULATION OF A SIGMA FACTOR  
; TITLE OF INVENTION: FROM MYCOBACTERIUM TUBERCULOSIS  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Pillsbury Madison & Sutro, L.L.P.  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/622.352A  
FILING DATE: 27-MAR-1996  
CLASSIFICATION: 435  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 242 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mycobacterium tuberculosis  
US-08-622-352A-3

Query Match 65.9%; Score 29; DB 2; Length 242;  
Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLEPDVAV 9  
Db 139 FEDLDFDAV 147

RESULT 31  
US-08-826-390-3  
Sequence 3, Application US/08826390  
Patent No. 6004764  
GENERAL INFORMATION:  
APPLICANT: Bishai, William R.  
APPLICANT: Young, Douglas B.  
APPLICANT: Zhang, Ying  
APPLICANT: Demajo, James  
TITLE OF INVENTION: Stationary Phase, Stress Response  
TITLE OF INVENTION: Sigma Factor From Mycobacterium tuberculosis, and  
TITLE OF INVENTION: Regulation Thereof  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/826.390  
FILING DATE: 27-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/622.353  
FILING DATE: 27-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/622.352  
FILING DATE: 27-MAR-1996  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 242 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mycobacterium tuberculosis  
US-08-826-390-3

Query Match 65.9%; Score 29; DB 3; Length 242;  
Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLEPDVAV 9  
Db 139 FEDLDFDAV 147

RESULT 32  
US-09-540-236-3467  
Sequence 3467, Application US/09540236  
Patent No. 6673910  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARR  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540.236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 3467  
LENGTH: 316  
TYPE: PRT  
ORGANISM: M.catarrhalis  
US-09-540-236-3467

Query Match 65.9%; Score 29; DB 4; Length 316;  
Best Local Similarity 75.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLEFDA 8  
Db 223 FLPAEFDA 230

RESULT 33  
US-09-270-767-44624  
Sequence 44624, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270.767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 44624  
LENGTH: 389  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-270-767-44624

Query Match 65.9%; Score 29; DB 4; Length 389;  
Best Local Similarity 75.0%; Pred. No. 4.5e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LQLEPDVAV 9  
Db 342 LQLEPDVAV 349

RESULT 34  
US-09-252-991A-19612  
Sequence 19612, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A

CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 19612  
LENGTH: 427  
TYPE: PRS  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19612

Query Match 65.9%; Score 29; DB 4; Length 427;  
Best Local Similarity 66.7%; Pred. No. 4.9e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
DB 247 FLDLQVDAV 255

RESULT 35  
US-09-107-532A-6186  
Sequence 6186, Application US/09107532A  
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354

COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 893-5007  
TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 6186:  
SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...467  
SEQUENCE DESCRIPTION: SEQ ID NO: 6186:

US-09-107-532A-6186

Query Match

65.9%; Score 29; DB 4; Length 467;

Best Local Similarity 55.6%; Pred. No. 5.4e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLOLEPDAV 9  
DB 304 FLDLQVDAV 312

RESULT 36  
US-09-107-532A-7248  
Sequence 7248, Application US/09107532A  
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354

COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 893-5007  
TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 7248:  
SEQUENCE CHARACTERISTICS:

LENGTH: 512 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...512  
SEQUENCE DESCRIPTION: SEQ ID NO: 7248:

US-09-107-532A-7248

Query Match

65.9%; Score 29; DB 4; Length 512;

Best Local Similarity 62.5%; Pred. No. 5.9e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDA 8  
DB 48 FLDLQVDAV 55

RESULT 37

US-09-252-991A-19386

Sequence 19386, Application US/09252991A  
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354

COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 893-5007  
TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 7248:  
SEQUENCE CHARACTERISTICS:

LENGTH: 512 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...512  
SEQUENCE DESCRIPTION: SEQ ID NO: 7248:

US-09-107-532A-7248

Query Match

65.9%; Score 29; DB 4; Length 512;

Best Local Similarity 62.5%; Pred. No. 5.9e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDA 8  
DB 48 FLDLQVDAV 55

RESULT 37

US-09-252-991A-19386

Sequence 19386, Application US/09252991A  
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354

COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 19386  
LENGTH: 642  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19386

Query Match 65.9%; Score 29; DB 4; Length 642;  
Best Local Similarity 87.5%; Pred. No. 7.5e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQLEFDAY 9  
DB 218 LQLEGDAY 225

RESULT 38  
US-08-674-351-2  
Sequence 2, Application US/08674351  
Patent No. 5831013  
GENERAL INFORMATION:  
APPLICANT: Bruem, Jeremy A.  
APPLICANT: Yao, Wensheng  
TITLE OF INVENTION: CAPSID POLYPEPTIDES AND USE TO INHIBIT  
TITLE OF INVENTION: VIRAL PACKAGING  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/674,351  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Timian, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 19226/740  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 263-1636  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 680 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-674-351-2

Query Match 65.9%; Score 29; DB 2; Length 680;  
Best Local Similarity 100.0%; Pred. No. 7.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LEFDAY 9

DB 434 LEFDAY 439

RESULT 39  
US-09-267-311-2  
Sequence 2, Application US/09267311  
Patent No. 6440715  
GENERAL INFORMATION:  
APPLICANT: XU, Shuang-yong  
TITLE OF INVENTION: Method for Cloning and Expression Of Rhodothermus  
TITLE OF INVENTION: Obamensis DNA Polymerase I Large Fragment In E. Coli  
FILE REFERENCE: NEB-157  
CURRENT APPLICATION NUMBER: US/09/267,311  
CURRENT FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
LENGTH: 924  
TYPE: PRT  
ORGANISM: Rhodothermus obamensis  
US-09-267-311-2

Query Match 65.9%; Score 29; DB 4; Length 924;  
Best Local Similarity 55.6%; Pred. No. 1.1e+03;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQLEFDAY 9  
DB 276 FQLEFDL 284

RESULT 40  
US-09-091-117-5  
Sequence 5, Application US/09091117  
Patent No. 6171589  
GENERAL INFORMATION:  
APPLICANT: The University of Melbourne  
TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and  
TITLE OF INVENTION: Vaccines  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: United States of America  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/091,117  
FILING DATE: 12 JUNE 1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU96/00803  
FILING DATE: 13-DEC-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PN7127  
FILING DATE: 13-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: WINNER, Ellen P.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: +1 303 499 8080  
TELEFAX: +1 303 499 8089  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1024 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein



ORIGINAL SOURCE:  
ORGANISM: Mycoplasma genitalium  
US-09-091-117-5

Query Match 65.9%; Score 29; DB 3; Length 1024;  
Best Local Similarity 66.7%; Pred. No. 1.2e+03;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLEPDV 9  
|||:|  
Db 465 FLOEQFDIV 473

Search completed: January 12, 2005, 20:17:39  
Job time : 23.6 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:37:43 ; Search time 201.4 Seconds  
(without alignments)  
16.031 Million cell updates/sec

Title: US-09-870-216c-5  
Perfect score: 44  
Sequence: 1 FLQLEFDAY 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues  
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04: \*  
1: geneseqp1980s: \*  
2: geneseqp1990s: \*  
3: geneseqp2000s: \*  
4: geneseqp2001s: \*  
5: geneseqp2002s: \*  
6: geneseqp2003as: \*  
7: geneseqp2003bs: \*  
8: geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	9	5 ABB08362	Abb08362 Synthetic
2	44	100.0	9	7 ABR82214	AbR82214 Human ant
3	44	100.0	352	5 ABB08367	Abb08367 Human ant
4	37	84.1	1278	4 AAB62030	Aab62030 Recombina
5	35	79.5	282	6 ADB06652	ADB06652 Allioiococ
6	35	79.5	285	6 ADB06654	ADB06654 Allioiococ
7	35	79.5	1612	4 ABB59769	Abb59769 Drosophila
8	33	75.0	217	7 ABBM74402	Abbm74402 DNA clone
9	33	75.0	374	8 ADP98863	Adp98863 C. albica
10	32	72.7	72	5 ABP06827	Abp06827 Human ORF
11	32	72.7	173	6 ABUT0578	Abu70578 Human adi
12	32	72.7	225	6 ABU21843	Abu21843 Protein e
13	32	72.7	274	6 ABU24075	Abu24075 Protein e
14	32	72.7	415	7 ADD27919	Add27919 Human PCO
15	32	72.7	419	7 ABO78318	AbO78318 Pseudomon
16	32	72.7	514	7 ADD12559	Add12559 Human ENZ
17	32	72.7	551	7 ADC24201	Adc24201 Human NOV
18	32	72.7	551	7 ADD12554	Add12554 Human ENZ
19	32	72.7	552	2 AAY29648	Aay29648 Sheep pro
20	32	72.7	552	7 ADD12555	Add12555 Human ENZ
21	32	72.7	562	5 ABB09701	Abb09701 Antino aci
22	32	72.7	580	7 ADD27944	Add27944 COX-1 ami
23	32	72.7	580	7 ADD27968	Add27968 COX-1 ami
24	32	72.7	599	2 AAR21690	Aar21690 Prostagla
25	32	72.7	599	5 ABG96403	Abg96403 Human ova

26	32	72.7	599	6 ABR42249	AbR42249 Human CYC
27	32	72.7	599	7 ADC24199	Adc24199 Human NOV
28	32	72.7	599	7 ADD27936	Add27936 Human COX
29	32	72.7	599	8 ADN05626	Adn05626 Antipepti
30	32	72.7	600	1 AAP91008	Aap91008 Prostagla
31	32	72.7	600	7 ADD27937	Add27937 Ovine COX
32	32	72.7	602	5 ABB57303	Abb57303 Mouse lsc
33	32	72.7	602	6 ABR42251	AbR42251 Mouse cyc
34	32	72.7	602	7 ADD27940	Add27940 Rat COX-1
35	32	72.7	602	7 ADD27940	Add27940 Murine CO
36	32	72.7	603	7 ADD27938	Add27938 Canine CO
37	32	72.7	606	7 ADD27939	Add27939 Rabbit CO
38	32	72.7	608	5 ABB07241	Abb07241 Canine cy
39	32	72.7	608	5 ABB07243	Abb07243 Human cyc
40	32	72.7	629	7 ADD27929	Add27929 Human hCO
41	32	72.7	633	5 ABG30579	Abg30579 Dog prost
42	32	72.7	633	8 ADOS2596	Ados2596 Dog COX-1
43	32	72.7	634	7 ADD27916	Add27916 Human COX
44	32	72.7	2289	2 AAW14987	Aaw14987 Protein d
45	31	70.5	129	4 AAU56952	Aau56952 Proponib

## ALIGNMENTS

RESULT 1	
ABBO8362	
ID	ABBO8362 standard; protein; 9 AA.
AC	ABBO8362;
XX	
DT	07-MAY-2002 (first entry)
XX	
DE	Synthetic epitope 2 of human cancer antigen eIF3.
XX	
KW	Human; melanoma antigen eukaryotic initiation factor 3; eIF3;
KW	ovarian cancer; MHC; cytostatic; immunomodulator; immune effector cell;
XX	anti-cancer; vaccine.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	Domain
FT	Location/Qualifiers
FT	1
FT	/note= "HLA-2 binding residue"
FT	2
FT	/note= "HLA-2 binding residue"
FT	3..8
FT	/note= "T-cell receptor (TCR) binding domain"
FT	9
FT	/note= "HLA-2 binding residue"
PN	WO200192307-A2.
XX	
PD	06-DEC-2001.
XX	
PF	30-MAY-2001; 2001WO-US017456.
XX	
PR	31-MAY-2000; 2000US-0209391P.
PR	17-AUG-2000; 2000US-0226258P.
PR	20-DEC-2000; 2000US-0257008P.
PA	(GENZ ) GENZYME CORP.
XX	
PI	Nicolette CA;
XX	
DR	WPI; 2002-139606/18.
DR	N-PSDB; ABA97213.
XX	
PT	New therapeutic compounds useful against human ovarian cancer, for
PT	modulating immune response in a subject, and for generating antibodies
PT	that specifically recognize and bind to these molecules.
XX	
PS	Claim 29; Page 59; 68pp; English.

XX The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterized by expression of antigen E1F3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents synthetic epitope 2 of human cancer antigen e1f3

XX  
 SQ Sequence 9 AA;  
 Query Match 100.0%; Score 44; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFDAY 9  
 |||||  
 Db 1 FLOLEFDAY 9

RESULT 2  
 ABR82214  
 ID ABR82214 standard; peptide; 9 AA.  
 AC ABR82214;  
 XX  
 DT 13-OCT-2003 (first entry)  
 XX  
 DE Human antigen e1f3 derived compound 2.  
 XX  
 KW Eukaryotic translation initiation factor 3; e1f3; neoplasia; cancer;  
 KW cyrostatic; gene therapy; human; antigen.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO2003050543-A1.  
 XX  
 PD 19-JUN-2003.  
 XX  
 PF 05-DEC-2001; 2001WO-US047997.  
 XX  
 PR 05-DEC-2001; 2001WO-US047997.  
 XX  
 PA (GENZ ) GENZYME CORP.  
 XX  
 PI Nicolette CA;  
 XX  
 DR WPI: 2003-532936/50.  
 DR N-PSDB; ACC85031.  
 XX  
 XX Aiding in the diagnosis of a neoplastic condition, useful for treating  
 PT cancer and related malignancies comprises determining the amount of  
 PT expression of an e1f3 protein in a test sample isolated from the cell or  
 PT tissue.  
 XX  
 PS Claim 12; Page 30; 77pp; English.

XX The invention relates to aiding in the diagnosis of a neoplastic  
 CC condition or susceptibility to a neoplastic condition of an animal cell  
 CC or tissue. The method involves determining the amount of expression of an  
 CC eukaryotic translation initiation factor 3 (e1f3) protein in a test  
 CC sample isolated from the cell or tissue, and diagnosing a neoplastic

CC condition or susceptibility to a neoplastic condition based on the amount  
 CC of expression of the e1f3 protein. The methods, compounds and kits are  
 CC useful in therapeutics, diagnostic and screening methods for human cancer  
 CC and related malignancies, e.g. ovarian, breast, lung, colon, prostate,  
 CC pancreatic or gastrointestinal cancer, or melanoma. Sequences ABR82213-16  
 CC represent compounds derived from the human antigen e1f3

XX  
 SQ Sequence 9 AA;  
 Query Match 100.0%; Score 44; DB 7; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFDAY 9  
 |||||  
 Db 1 FLOLEFDAY 9

RESULT 3  
 ABB08367  
 ID ABB08367 standard; protein; 352 AA.  
 AC ABB08367;  
 XX  
 DT 07-MAY-2002 (first entry)  
 XX  
 DE Human cancer antigen e1f3 variant 2 amino acid sequence.  
 XX  
 DE Human; melanoma antigen eukaryotic initiation factor 3; e1f3;  
 KW ovarian cancer; MHC; cyrostatic; immunomodulator; immune effector cell;  
 KW anti-cancer; vaccine.  
 XX  
 OS Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT Misc-difference 242  
 FT /note= "wild-type Asn is replaced by Phe"  
 FT Domain 242  
 FT /note= "HLA-2 binding residue"  
 FT Domain 243  
 FT /note= "HLA-2 binding residue"  
 FT Domain 244..249  
 FT /note= "T-cell receptor (TCR) binding domain"  
 FT Misc-difference 246  
 FT /note= "wild-type Leu is replaced by Glu"  
 FT Misc-difference 247  
 FT /note= "wild-type Met is replaced by Phe"  
 FT Misc-difference 249  
 FT /note= "wild-type Arg is replaced by Ala"  
 FT Domain 250  
 FT /note= "HLA-2 binding residue"  
 PN WO200192307-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US017456.  
 XX  
 PR 31-MAY-2000; 2000US-0209391P.  
 PR 17-AUG-2000; 2000US-0226258P.  
 PR 20-DEC-2000; 2000US-0257008P.  
 XX  
 XX (GENZ ) GENZYME CORP.  
 PA  
 PI Nicolette CA;  
 XX  
 DR WPI; 2002-139606/18.  
 XX  
 XX New therapeutic compounds useful against human ovarian cancer, for  
 PT modulating immune response in a subject, and for generating antibodies  
 PT that specifically recognize and bind to these molecules.  
 XX  
 PS Claim 7; Page; 68pp; English.

XX The invention relates to novel therapeutic compounds, that are designed  
CC to enhance binding to MHC molecules and to enhance immunoregulatory  
CC properties relative to their natural counterparts. The activity of the  
CC compounds of the invention may be described as cytostatic and  
CC immunomodulatory. The compounds are useful against human ovarian cancer,  
CC for modulating immune response in a subject, and for generating  
CC antibodies that specifically recognize and bind to these molecules.  
CC Compositions comprising the compounds are useful as components of anti-  
CC cancer vaccines and to expand immune effector cells that are specific for  
CC cells characterized by expression of antigen E1F3 (melanoma antigen  
CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
CC a detectable agent may be used in diagnostic procedures, such as in the  
CC detection and purification of antibodies, and as immunogens for  
CC production of antibodies. The polynucleotides can be used as primers for  
CC detecting genes or gene transcripts expressed in APC to confirm  
CC transduction of the polynucleotides into host cells. The current sequence  
CC represents the human cancer antigen e1f3 variant 2 amino acid sequence.  
CC Note: This sequence is not present in the specification, but may be  
CC created from the sequence of the wild-type human cancer antigen e1f3  
CC sequence given in ABB08360  
XX

SQ Sequence 352 AA;

Query Match 100.0%; Score 44; DB 5; Length 352;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 F1QLPEPDV 9  
Db 242 F1QLPEPDV 250

RESULT 4

ID AAB62030 standard; protein; 1278 AA.

XX AAB62030;

DT 14-MAY-2001 (first entry)

DE Recombinant P. furiosus helicase 7.

XX MCM; minichromosome maintenance protein; archaeal polypeptide; PCNA;  
XX RFC-P38; RFC-P35; RFA; CDC6; FEN-1; dUTPase; ligase; helicase dna2; PCR;  
XX helicase 7; nucleic acid amplification; polymerase chain reaction.  
XX

OS Pyrococcus furiosus.

XX WO200109347-A2.

XX 08-FEB-2001.

PD 28-JUL-2000; 2000MO-US020532.

PF 30-JUL-1999; 99US-0146580P.

XX (STRA-) STRATAGENE.

XX Hogrefe HH, Cline JM, Hansen CJ, Borns MC;

XX WPI; 2001-182959/18.

DR N-PSDB; AAF57034.

XX Composition for improving nucleic acid polymerase reactions, useful e.g.  
XX in synthesis or amplification, contains at least one archaeal accessory  
XX protein.

XX Claim 182; Fig 34; 147pp; English.

XX The invention provides a composition (A) for enhancing nucleic acid  
XX polymerase reactions that comprises an archaeal MCM (minichromosome  
XX maintenance protein) and at least one of the archaeal polypeptides (PCNA,

CC RFC-P38 or -P35, RFA, CDC6, FEN-1, dUTPase, ligase, helicase dna2, or  
CC helicases 2-8). (A) And similar compositions containing different  
CC combinations of accessory proteins, are used to improve performance of  
CC synthesis, amplification, mutagenizing, labeling and detecting reactions,  
CC e.g. for gene characterization, cloning, detection of allelic variants,  
CC diagnosis and screening for disease, particularly where done by  
CC polymerase chain reaction (PCR). Some of the proteins also stabilize  
CC duplexes during polymerase reactions or improve exonuclease reactions,  
CC for example RFA also improves specificity of nucleic acid/protein  
CC interaction and PCNA improves polymerase-mediated repair processes and  
CC hybridization reactions. Nucleic acids encoding the archaeal polypeptides  
CC are used for recombinant production of proteins, and fragments of the  
CC nucleic acid as probes and primers for screening related sequences. The  
CC accessory proteins increase accuracy and efficiency of polymerase  
CC reactions, allow use of lower denaturation and extension temperatures  
CC (possibly isothermal processing), and improve synthesis of long targets.  
CC The present sequence represents a P. furiosus recombinant helicase 7  
XX

SQ Sequence 1278 AA;

Query Match 84.1%; Score 37; DB 4; Length 1278;  
Best Local Similarity 87.5%; Pred. No. 1.4e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 F1QLPEFDA 8  
Db 650 Y1QLPEFDA 657

RESULT 5  
ID ADB06652 standard; protein; 282 AA.  
XX ADB06652;

DT 20-NOV-2003 (first entry)

DE Alloiococcus ostitis antigenic protein SEQ ID NO:592.

XX Alloiococcus ostitis; antigenic protein; immunogenic; immunisation;  
XX gene therapy; Gram-positive bacterium; infection.  
XX

OS Alloiococcus ostitis.

XX WO2003048304-A2.

PD 12-JUN-2003.

PF 25-NOV-2002; 2002MO-US036123.

XX 29-NOV-2001; 2001US-0333777P.

PR 18-NOV-2002; 2002US-0426742P.

XX (AMHP) WYETH HOLDINGS CORP.

XX Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;

XX WPI; 2003-505284/47.

DR N-PSDB; ADB06651.

XX New Alloiococcus ostitis polynucleotides and polypeptides, useful for  
XX treating and diagnosing diseases, drug screening assays and monitoring of  
XX effects during drug clinical trials.

PS Claim 33; SEQ ID NO 592; 1019pp; English.

XX The present invention describes an isolated polynucleotide (1) of  
XX Alloiococcus ostitis genomic DNA, which encodes an antigenic protein.  
XX Alloiococcus ostitis is a Gram-positive bacterium. Also described: (1)  
XX an isolated polypeptide that is encoded by the polynucleotide (1); (2) an  
XX expression vector comprising the novel isolated polynucleotide (1), its  
XX complement, degenerate variant or fragment; (3) a genetically engineered  
XX host cell, transfected, transformed or infected with the vector of (2);

(4) an antibody specific for the polypeptide of (1); (5) an immunogenic composition comprising the polypeptide, its complement, biological equivalent or fragment, or the polynucleotide that is comprised in the expression vector; (6) a pharmaceutical composition comprising the polypeptide of (1) and a carrier; (7) a protein chip comprising an array of the polypeptides of (1), their biological equivalent or fragment; (8) immunising against *Allotococcus otitidis* by administering to a host the immunogenic composition; (9) detecting and/or identifying *Allotococcus otitidis* in the biological sample; (10) a kit comprising a container containing the novel polynucleotide, its degenerate variant or fragment, or the antibody of (4); and (11) producing a polypeptide by culturing the genetically engineered host cell under conditions suitable to produce the polypeptide from the culture. (1) can be used in gene therapy. The polynucleotides, polypeptides, antibodies and compositions of the present invention can be used for treating and diagnosing diseases, drug screening assays and monitoring of effects during drug clinical trials. The polynucleotides are useful for expressing and detecting *Allotococcus otitidis*. The present sequence represents an *Allotococcus otitidis* antigen protein from the present invention.

CC Sequence 282 AA;

Query Match 79.5%; Score 35; DB 6; Length 282;  
Best Local Similarity 77.8%; Pred. No. 72;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEFPAV 9  
| | | | | | | | | |  
Db 184 FLOLEIDSV 192

RESULT 6  
ADB06654  
ID ADB06654 standard; protein; 285 AA.

XX ADB06654;

XX 20-NOV-2003 (first entry)

XX *Allotococcus otitidis* antigenic protein SEQ ID NO:594.

XX *Allotococcus otitidis*; antigenic protein; immunogenic; immunisation;  
XX gene therapy; Gram-positive bacterium; infection.

XX *Allotococcus otitidis*.

XX WO2003048304-A2.

XX 12-JUN-2003.

XX 25-NOV-2002; 2002WO-US036123.

XX 29-NOV-2001; 2001US-0333777P.

XX 18-NOV-2002; 2002US-0426742P.

XX (AMHP ) WYETH HOLDINGS CORP.

XX Fletcher ID, McMichael JC, Russell DP, Zagursky RJ;

XX WPI; 2003-505284/47.

XX N-PSDB; ADB06653.

XX New *Allotococcus otitidis* polynucleotides and polypeptides, useful for  
XX treating and diagnosing diseases, drug screening assays and monitoring of  
XX effects during drug clinical trials.

XX Claim 33; SEQ ID NO 594; 1019PP; English.

XX The present invention describes an isolated polynucleotide (1) of  
XX *Allotococcus otitidis* genomic DNA, which encodes an antigenic protein.  
XX *Allotococcus otitidis* is a Gram-positive bacterium. Also described: (1)  
XX an isolated polypeptide that is encoded by the polynucleotide (1); (2) an  
XX expression vector comprising the novel isolated polynucleotide (1); its

CC complement, degenerate variant or fragment; (3) a genetically engineered  
CC host cell, transfected, transformed or infected with the vector of (2);  
CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic  
CC composition comprising the polypeptide, its complement, biological  
CC equivalent or fragment, or the polynucleotide that is comprised in the  
CC expression vector; (6) a pharmaceutical composition comprising the  
CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array  
CC of the polypeptides of (1), their biological equivalent or fragment; (8)  
CC immunising against *Allotococcus otitidis* by administering to a host the  
CC immunogenic composition; (9) detecting and/or identifying *Allotococcus otitidis* in the biological sample; (10) a kit comprising a container containing the novel polynucleotide, its degenerate variant or fragment, or the antibody of (4); and (11) producing a polypeptide by culturing the genetically engineered host cell under conditions suitable to produce the polypeptide from the culture. (1) can be used in gene therapy. The polynucleotides, polypeptides, antibodies and compositions of the present invention can be used for treating and diagnosing diseases, drug screening assays and monitoring of effects during drug clinical trials. The polynucleotides are useful for expressing and detecting *Allotococcus otitidis*. The present sequence represents an *Allotococcus otitidis* antigen protein from the present invention.

CC Sequence 285 AA;

Query Match 79.5%; Score 35; DB 6; Length 285;  
Best Local Similarity 77.8%; Pred. No. 73;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEFPAV 9  
| | | | | | | | | |  
Db 187 FLOLEIDSV 195

RESULT 7  
ABB59769  
ID ABB59769 standard; protein; 1612 AA.

XX ABB59769;

XX 26-MAR-2002 (first entry)

XX *Drosophila melanogaster* polypeptide SEQ ID NO 6099.

XX *Drosophila*; developmental biology; cell signalling; insecticide;  
XX pharmaceutical.

XX *Drosophila melanogaster*.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL03872.

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from *Drosophila* and for elucidating cell signaling and cell-cell  
XX interactions.

XX Disclosure; SEQ ID NO 6099; 21PP + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from *Drosophila*. The invention is  
XX useful in developmental biology and in elucidating cell signalling and

cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (AAB57737-CC ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO a/fcp.wipo.int/pub/published\_pat\_sequences

**SQ Sequence 1612 AA;**

Query Match	79.5%	Score 35;	DB 4;	Length 1612;
Best Local Similarity	66.7%	Pred. No. 4.6e+02;		
Matches	6;	Conservative	3;	Mismatches 0;
			Indels	0;
			Gaps	0;

```

QY      1 FLOLEFPAV 9
          :|:|||||:
DB      254 YLELEFDAL 262

```

RESULT 8  
ABM74402  
ID ABM74402 standard; protein; 217 AA.

DT 17-OCT-2003 (first entry)

DNA clone originating in barley containing SNP sequence #812.

KW Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis.

05 *Hordeum vulgare.*  
xy

PN WO2003057877-A1.

PD 17-JUL-2003

PF 16-DEC-2002; 2002WO-IB005403.  
yy

PR 20-DEC-2001; 2001JP-00387059.  
DE 20 DEC 2001 2001TD 00387131

PR 20-DEC-2001; 2001JP-004032299.  
20 DEC 2001 004032200

PR 27-SEP-2002; 2002JP-00327515.  
xx

PA (UYN1-) UNIV JAPAN OKAYAMA.  
xy

PI Sato K, Takeda K, Kohara Y;  
xx

DR WPI; 2003-587127/55.

PT Single nucleotide polymorphism sites in barley varieties and DNA sequences containing them for analysis and identification of barley varieties and production of barley transformants with desired characteristics.

PS Disclosure; SEQ ID XX; 284pp; Japanese.

CC The present invention relates to oligonucleotide clones originating in  
CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms  
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley  
CC varieties, identification of particular varieties and genotype-phenotype  
CC analysis, isolation of specific genes and creation of new varieties by  
CC transformation of barley varieties with them and production of new barley  
CC varieties with desired properties. The present sequence represents an  
CC oligonucleotide clone sequence featured in the specification. The  
CC sequence data for this patent did not form part of the primed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp.wipo.int/pub/published-pct-sequences](http://wipo.int/pub/published-pct-sequences)

**SQ** Sequence 217 AA;

Query Match 75.0%; Score 33; DB 7; Length 217;

Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFDAY 9  
|||:::|  
Db 44 FLOIKFNAY 52

RESULT 9  
ADP98863  
ID ADP98863 standard; protein; 374 AA

AC ADP98863 ;  
XY

DT 23-SEP-2004 (first entry)  
 VV

DE C. albicans specific gene, orf6.4105, protein sequence.

KM Diploid fungal cell; allele; gene disruption cassette;  
KM promoter replacement fragment; antifungal; fungicide; gene therapy;  
KM infection; *Candida albicans*.

*Candida albicans*.

PN WO2004056965-A2.

PD 08-JUL-2004

PF 19-DEC-2003; 2003WO-US040618.

PR 19-DEC-2002; 2002US-0434832P.

PA (ELIT-) ELITRA PHARM INC.  
PA (ELIT-) ELITRA PHARM INC.

PA (ELIT-) ELITRA CANADA LTD  
...

PI Roemer T, Jiang B, Boone C, Bussey H, ...

DR WPI; 2004-500296/47.

25

PT Constructing a strain of diploid fungal cells in which both alleles of a gene are modified comprises modifying the alleles of a gene in the fungus by recombination using a gene disruption cassette and a promoter replacement fragment.

PS Claim 44; SEQ ID NO 7038; 163pp; English.

The invention relates to a novel method for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The method comprises modifying the alleles of a gene in diploid fungal cells by recombination using a gene disruption cassette and a promoter replacement fragment. The invention further comprises: assembling a collection of diploid fungal cells each of which comprises modified alleles of a different gene; a strain of diploid fungal cells comprising modified alleles of a gene, where the first allele of the gene is inactivated by a gene disruption cassette comprising a nucleotide sequence encoding an expressible selectable marker; and the expression of the second allele of the gene is regulated by a heterologous promoter that is operably linked to the coding region of the second allele of the gene; and where the gene encodes the polypeptide mentioned above; a collection of diploid fungal strains comprising the diploid strains cited above, where substantially all the different genes that encode the above amino acid sequences are modified and are present in different diploid strains in the collection; a nucleic acid molecule microarray comprising nucleic acid molecules, where each nucleic acid molecule comprises a nucleotide sequence that is hybridizable to a target nucleotide sequence comprising any of the 310 nucleotide sequences listed in the specification (ADP98316-ADP98257), identifying a gene that is essential to the survival or growth of a fungus, that contributes to the resistance and/or pathogenicity of a fungus, or that contributes to the resistance of a diploid fungus to an antifungal agent; identifying an antifungal agent that inhibits the growth of a diploid fungus, or a therapeutic agent for treatment of a mammalian disease; correlating changes in the

levels of proteins or gene transcripts with the inhibition of growth or proliferation of a diploid fungal cell; a purified or isolated nucleic acid molecule comprising a nucleotide sequence encoding a gene product required for proliferation of *Candida albicans*, where the gene product consists of any of the above-mentioned amino acid sequences; a vector comprising a promoter operably linked to the nucleic acid molecule cited above; a host cell containing the vector; a purified or isolated polypeptide comprising any of the 61 amino acid sequences given in the specification (ADP96718-ADP96778); a fusion protein comprising a fragment of a first polypeptide fused to a second polypeptide, the fragment consisting of at least 6 consecutive residues of any of ADP98826-ADP99135; producing a polypeptide; identifying a compound which modulates the activity of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825, eliciting an immune response in an animal; a strain of *Candida albicans*, where a first allele of a gene comprising any of ADP98516-ADP98825 is inactive and a second allele of the gene is under the control of a heterologous promoter; identifying a compound or binding partner that binds to the polypeptide comprising any of ADP98826-ADP99135, or its fragment; identifying a compound having the ability to inhibit growth or proliferation of *Candida albicans*; inhibiting growth or proliferation of *Candida albicans* cells; manufacturing an antimycotic compound; treating an infection of a subject by *Candida albicans*; preventing or containing contamination of an object by *Candida albicans*, or for preventing or inhibiting formation on a surface of a biofilm comprising *Candida albicans*; a pharmaceutical composition comprising a therapeutic amount of an agent which reduces the activity or level of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825 in a pharmaceutical carrier; an antibody preparation which binds the polypeptide; methods for evaluating a compound against a target gene product encoded by any of ADP98516-ADP98825; identifying an antimycotic compound; a computer or a computer readable medium that comprises at least one of the nucleotide sequences mentioned in the specification or at least one amino acid sequence selected from ADP98826-ADP99135; a method assisted by a computer for identifying a putatively essential gene of a fungus; and a protein array comprising proteins, where at least one protein comprises an amino acid sequence or a portion of an amino acid sequence selected from ADP98516-ADP98825. The novel methods and compositions have fungicidal activity. The compositions may be used in gene therapy. The composition and methods are useful for drug screening purposes or for diagnosing, preventing or treating infections associated with *Candida albicans*. These may also be used for constructing strains useful for identification and validation of gene products as effective targets for therapeutic intervention, for identifying and validating gene products as effective targets for therapeutic intervention, and for collecting identified essential genes. This sequence represents the protein of a *Candida albicans* fungal specific gene of the invention.

NOTES: This sequence was downloaded from an electronic sequence listing provided on the WIPO website.

```

Query Match 75.0%; Score 33; DB 8; length 374;
Best Local Similarity 85.7%; Pred No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 QLEPDPAV 9
    |||||:
Db 293 QLEPDPAI 299

RESULT 10
ABP06827
ID ABP06827 standard; protein; 72 AA.
XX
AC ABP06827;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX protein sequence SEQ ID NO:13636.
XX

Human; open reading frame; ORFX; gene therapy; cancer; cirtrosis;
KM hyperproliferative disorder; psoriasis; benign tumor; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;

```

KV	cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KV	hyperension; hypothyroidism; cholesterol ester storage disease;
KW	immune deficiency; immune disorder; infectious disease;
KW	autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW	myasthenia gravis.
XX	
OS	Homo sapiens.
XX	
PN	WO200192523-A2.
XX	
PD	06-DEC-2001.
XX	
PF	29-MAY-2001; 2001WO-US010836.
XX	
PR	30-MAY-2000; 2000US-0206132P.
PR	29-AUG-2000; 2000US-0228716P.
XX	
XX	
PA	(CURA-) CURAGEN CORP.
XX	
PI	Shimkets RA, Leach MD;
XX	
DR	WPI; 2002-106308/14.
DR	N-PSDB; AEN22579.
XX	
XX	
PT	Novel human polypeptides and polymucleotides useful for diagnosing,
PT	preventing and treating cardiovascular disease, neurodegenerative,
PT	hyperproliferative disorders and autoimmune disorders.
XX	
PS	Disclosure; SEQ ID NO 13636; 1037bp; English.

Query Match	Best Local Similarity	Score 32;	DB 5;	Length 72;
Matches 6;	Conservative	2;	Mismatches 1;	Indels 0;
<p>           1 FLOLEPDVAV 9                 :            54 FLQGLMDVDV 62         </p>				
<p>           RESULT 11            ABU70578            ID ABU70578 standard; protein; 173 AA.            AC ABU70578;            XX            DT 10-JUN-2003 (first entry)         </p>				



XX	Human adipocyte Selected Interacting domain, SID, #209.
DE	
XX	
KM	Human; prey; adipocyte; SID; selected interacting domain; anorectic;
KM	antidiabetic; protein-protein interaction; diabetes;
KM	yeast 2-hybrid assay; metabolic disorder; obesity.
XX	
OS	Homo sapiens.
XX	
PN	WO200286122-A2.
XX	
PD	31-OCT-2002.
XX	
PF	14-MAR-2002; 2002WO-EP003768.
PR	14-MAR-2001; 2001US-0275734P.
XX	
PA	(HYBR-) HYBRIGENICS.
XX	
PI	Legrain P, Davlet L;
XX	
DR	WPI; 2003-103412/09.
DR	N-PSDB; ACA571122.
XX	
PT	New complex between two interacting proteins in adipocyte cells, useful
PT	for identifying selected interacting domains that modulate protein
PT	interactions, or for preventing or treating metabolic disorders such as
XX	obesity or diabetes.
XX	
PS	Claim 6; Page 175; 382pp; English.
XX	
CC	The invention relates to a complex between two interacting proteins in
CC	adipocyte cells, given in the specification. The proteins are identified
CC	by selecting a bait protein from a known adipocyte marker and then
CC	performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC	members of an adipocyte cDNA library. The proteins are designated SID
CC	(RTM) (selected interacting domains) proteins. Also included are a
CC	polynucleotide encoding a polypeptide in the adipocyte cells, a
CC	recombinant host cell expressing at least one of the interacting
CC	polypeptides of the complex, selecting a modulating compound in adipocyte
CC	cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
CC	sequences given in the specification (including its fragment or variant),
CC	a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC	given in the specification (including its fragment or variant), a vector
CC	comprising the SID (RTM) polynucleotide, a recombinant host cell
CC	comprising the vector, a protein chip comprising the polypeptides and a
CC	record comprising all or part of the data, listed in the specification.
CC	The complex, polypeptides, polynucleotides and compounds are useful for
CC	preventing or treating metabolic disorders such as obesity or diabetes.
CC	The polynucleotides are useful as probes or primers. The complex is
CC	particularly useful for identifying selected interacting domains (SID
CC	(RTM)) for screening drugs that modulate the protein interaction, thus
CC	exhibiting the therapeutic effect. The present sequence represents a SID
CC	(prey) protein of the invention
XX	
XX	
SQ	Sequence 173 AA;
QY	
Db	
	1 FLQLQKPD 7
	141 FLQLKPD 147
RESULT 12	
ABU21843	72.7%; Score 32; DB 6; Length 173;
ID	ABU21843 standard; protein; 225 AA.
XX	
AC	ABU21843;
XX	
PT	19-JUN-2003 (first entry)

XX	Protein encoded by Prokaryotic essential gene #7370.
XX	
KW	Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	
OS	Burkholderia fungorum.
XX	
PN	MO200277183-A2.
PD	
PD	03-OCT-2002.
PF	21-MAR-2002; 2002WO-US009107.
XX	
XX	21-MAR-2001; 2001US-00815242.
PR	06-SEP-2001; 2001US-00948993.
PR	25-OCT-2001; 2001US-0342923P.
PR	08-FEB-2002; 2002US-00072851.
PR	06-MAR-2002; 2002US-0362639P.
XX	
PA	(BLIT-) BLITRA PHARM INC.
XX	
P1	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KT, Zykkind JW;
P1	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX	
DR	WPI; 2003-029926/02.
XX	
XX	N-PSDB; ACA25713.
PT	New antisense nucleic acids, useful for identifying proteins or screening
PT	for homologous nucleic acids required for cellular proliferation to
PT	isolate candidate molecules for rational drug discovery programs.
XX	
XX	Claim 25; SEQ ID NO 49767; 1766pp; English.
CC	
CC	The invention relates to an isolated nucleic acid comprising any one of
CC	the 6213 antisense sequences given in the specification where expression
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:
CC	(1) a vector comprising a promoter operably linked to the nucleic acid
CC	encoding a polypeptide whose expression is inhibited by the antisense
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC	polypeptide or its fragment whose expression is inhibited by the
CC	antisense nucleic acid; (4) an antibody capable of specifically binding
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC	proliferation or the activity of a gene in an operon required for
CC	proliferation; (7) identifying a compound that influences the activity of
CC	the gene product or that has an activity against a biological pathway
CC	required for proliferation, or that inhibits cellular proliferation; (8)
CC	identifying a gene required for cellular proliferation or the biological
CC	pathway in which a proliferation-required gene or its gene product lies
CC	or a gene on which the test compound that inhibits proliferation of an
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC	compound's activity; (11) a culture comprising strains in which the gene
CC	product is overexpressed or underexpressed; (12) determining the extent
CC	to which each of the strains is present in a culture or collection of
CC	strains; or (13) identifying the target of a compound that inhibits the
CC	proliferation of an organism. The antisense nucleic acids are useful for
CC	identifying proteins or screening for homologous nucleic acids required
CC	for cellular proliferation to isolate candidate molecules for rational
CC	drug discovery programs, or for screening homologous nucleic acids
CC	required for proliferation in cells other than <i>S. aureus</i> , <i>S. typhimurium</i> ,
CC	<i>K. pneumoniae</i> or <i>P. aeruginosa</i> . The present sequence is encoded by one of
CC	the target prokaryotic essential genes. Note: The sequence data for this
CC	patent did not form part of the printed specification, but was obtained
CC	in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 225 AA;
XX	
Query Match	72.7%; Score 32; DB 6; Length 225;
Best Local Similarity	66.7%; Pred. No. 2,3+02;
Matches	6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XY	1 FLOLEFPAV 9

Db 111 YLKLEDPDV 119

RESULT 13  
ABU24075  
ID ABU24075 standard; protein; 274 AA.  
XX  
XX ABU24075;  
XX  
XX 19-JUN-2003 (first entry)  
XX  
XX Protein encoded by Prokaryotic essential gene #9602.  
DE  
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
XX Clostridium acetobutylicum.  
XX  
XX WO200277183-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 21-MAR-2002; 2002WO-US009107.  
XX  
XX 21-MAR-2001; 2001US-00815242.  
XX  
XX 06-SEP-2001; 2001US-00948993.  
XX  
XX 25-OCT-2001; 2001US-0342923P.  
XX  
XX 08-FEB-2002; 2002US-00072851.  
XX  
XX 06-MAR-2002; 2002US-0362699P.  
XX  
XX (ELITR-) ELITRA PHARM INC.  
XX  
XX Wang L, Zammudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,  
XX  
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
XX WPI; 2003-029926/02.  
XX  
XX N-PSDB; ACA27945.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
XX  
XX for homologous nucleic acids required for cellular proliferation to  
XX  
XX isolate candidate molecules for rational drug discovery programs.  
XX  
XX  
XX Claim 25; SEQ ID NO 51999; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway (8)  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX SQ Sequence 274 AA;  
XX  
XX Query Match 72.7%; Score 32; DB 6; Length 274;  
XX  
XX Best Local Similarity 55.6%; Pred. No. 2.9e+02;  
XX  
XX Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX QY 1 FLQLEPDV 9  
XX  
XX DB 205 FLKLDPDGI 213  
XX  
XX  
XX RESULT 14  
XX  
XX ADD27919  
XX  
XX ID ADD27919 standard; protein; 415 AA.  
XX  
XX  
XX ADD27919;  
XX  
XX 15-JAN-2004 (first entry)  
XX  
XX  
XX Human PCOX-1a protein SEQ ID NO:5.  
XX  
XX  
XX  
XX Cyclooxygenase type 1; cyclooxygenase type 1 variant protein;  
XX  
XX COX-1 variant protein; genetic disease; tissue typing;  
XX  
XX forensic identification; human; PCOX-1a.  
XX  
XX Homo sapiens.  
XX  
XX PN WO2003029411-A2.  
XX  
XX PD 10-APR-2003.  
XX  
XX PF 28-SEP-2002; 2002WO-US030947.  
XX  
XX PR 28-SEP-2001; 2001US-0326133P.  
XX  
XX PR 15-APR-2002; 2002US-0373225P.  
XX  
XX PR 16-APR-2002; 2002US-0373661P.  
XX  
XX PR 16-SEP-2002; 2002US-0411575P.  
XX  
XX PA (UYYO ) UNIV BRIGHAM YOUNG.  
XX  
XX PI Simmons D, Chandrasekharan VN;  
XX  
XX DR WPI; 2003-421222/39.  
XX  
XX N-PSDB; ADD27918.  
XX  
XX PT Novel cyclooxygenase type 1 variant polypeptide and a polynucleotide  
XX  
XX encoding the polypeptide, useful for identifying a compound that binds to  
XX  
XX and modulates the activity of COX-1 variant polypeptide.  
XX  
XX PS Claim 18; SEQ ID NO 5; 150pp; English.

CC The present invention describes an isolated cyclooxygenase type 1 (COX-1)  
CC variant polypeptide (I). (I) is useful for identifying a compound which  
CC modulates the activity of (I). A nucleotide sequence encoding (I) can be  
CC used for mapping their respective genes on a chromosome, and so locating  
CC gene regions associated with genetic disease, identifying an individual  
CC from a minute biological sample (tissue typing), and to aid in forensic  
CC identification of a biological sample. The present sequence represents  
CC PCOX-1a, which is used in the exemplification of the present invention.

XX  
XX SQ Sequence 415 AA;  
XX  
XX Query Match 72.7%; Score 32; DB 7; Length 415;  
XX  
XX Best Local Similarity 85.7%; Pred. No. 4.5e+02;  
XX  
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX QY 1 FLQLEPD 7  
XX  
XX DB 170 FLQLEPD 176

RESULT 15	
ABO78318	
ID	ABO78318 standard; protein, 419 AA.
XX	
AC	ABO78318;
DT	29-JUL-2004 (first entry)
XX	
DE	Pseudomonas aeruginosa polypeptide #10493.
XX	
XX	Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
OS	Pseudomonas aeruginosa.
PN	US6551795-B1.
XX	
PD	22-APR-2003.
XX	
PF	18-FEB-1999; 99US-00252991.
PR	18-FEB-1998; 98US-0074788P.
PR	27-JUL-1998; 98US-0094190P.
XX	
PA	(GENO-) GENOME THERAPEUTICS CORP.
XX	
PI	Rubinfeld MJ, Nolling J, Deloughery C, Bush D;
XX	WPI; 2003-615309/58.
DR	N-PSDB; ABD11889.
PT	Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, of
PT	useful as molecular targets for diagnostics, prophylaxis and treatment of
XX	pathological conditions resulting from bacterial infection.
XX	
PS	Disclosure; SEQ ID NO 27064; 455pp; English.
XX	
CC	The invention relates to Pseudomonas aeruginosa polypeptides and the
CC	polynucleotides encoding them. The sequences are useful in diagnosis and
CC	therapy of pathological conditions, as molecular targets for diagnostics,
CC	prophylaxis and treatment of pathological conditions resulting from a
CC	bacterial infection, for evaluating a compound, such as a polypeptide,
CC	for the ability to bind a P. aeruginosa nucleic acid, as components of
CC	effective antibacterial targets, as targets for antibacterial drugs,
CC	including anti-P. aeruginosa drugs, as templates for recombinant
CC	production of P. aeruginosa-derived peptides or polypeptides, as target
CC	components for P. aeruginosa and/or treatment of P. aeruginosa-caused
CC	infection, and in detection of P. aeruginosa sequences or other sequences
CC	of Pseudomonas species using biopchip technology. Sequences ABO67826-
CC	ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC	sequence data for this patent did not form part of the printed
CC	specification but was obtained in electronic format from USPTO at
CC	seqdata.uspto.gov/sequence.html
XX	
XX	
SEQ	Sequence 419 AA;
QY	
QY	2 LQLEPDAY 9
DB	226 LQLEPDAL 233
XX	
RESULT 16	
ADD12559	
ID	ADD12559 standard; protein, 514 AA.
XX	
AC	ADD12559;
XX	
DT	01-JAN-2004 (first entry)
XX	
DE	Human ENZW-19 protein SEQ ID NO:19.

XX human; enzyme; ENZM; cytostatic; antiarteriosclerotic; anti-HIV;  
KW antiallergic; cerebroprotective; antiparkinsonian; anticonvulsant;  
KW neurotropic; neuroprotective; antiinflammatory; ophthalmological;  
KW antithyroid; antidiabetic; antibacterial; vincristide; protozoacide;  
KW antiparasitic; fungicide; anorectic; cardiant; hypotensive;  
KW antiferility; hepatotropic; gene therapy; autoimmune disorder;  
KW inflammatory disorder; AIDS; allergy; atopic dermatitis; arthritis;  
KW thyroiditis; infection; metabolic disorder; obesity;  
KW reproductive disorder; infertility; neurological disorder;  
KW Parkinson's disease; Alzheimer's disease; stroke; epilepsy;  
KW cardiovascular disorder; myocardial infarction; hypertension;  
KW eye disorder; cell proliferative disease; cancer; atherosclerosis;  
KW hepatitis.

OS Homo sapiens.  
XX  
XX  
XX WO2003072729-A2.  
XX  
XX  
XX 04-SEP-2003.  
XX  
XX 21-FEB-2003; 2003WO-US005478.  
XX  
XX 22-FEB-2002; 2002US-0359513P.  
XX 19-MAR-2002; 2002US-0365795P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Sprague WM, Kable AE, Lee S, Griffin JA, Becha SD, Lee EA;  
PI Hattalia AJA, Khare R, Emerling BM, Marguis JP, Ramkumar J;  
PI Elliott VS, Richardson TW, Baughn MW, Jin P, Chien D, Hawkins PR;  
PI Wilson AD, Chavala NK, Tran UK, Lee SY, Zebairadian Y, Jiang X;  
PI Jackson AA, Bhatia UC, Burrill JD, Blake JD, Ho A, Zheng W;  
XX  
XX WPI: 2003-731608/69.  
DR N-PSDB; ADD12616.  
XX  
XX  
XX New human enzymes and polynucleotides, useful for diagnosing, preventing  
PT or treating diseases or conditions associated with aberrant enzyme  
PT expression, e.g. cancer, hepatitis, AIDS, atherosclerosis, obesity,  
PT infection or stroke.  
XX  
XX  
XX Claim 1; SEQ ID NO 19; 426bp; English.

XX  
XX The present sequence represents a human enzyme designated ENZM-19. Human  
CC enzymes of the present invention have cytostatic, antiarteriosclerotic,  
CC anti-HIV, antiallergic, cerebroprotective, antiparkinsonian,  
CC anticonvulsant, neurotropic, neuroprotective, antiinflammatory,  
CC ophthalmological, antithyroid, antiarthritic, antibacterial, vincristide,  
CC protozoacide, antiparasitic, fungicide, anorectic, cardiant, hypotensive,  
CC antiferility and hepatotropic activities, and can be used in gene  
CC therapy. The human enzymes and polynucleotides encoding them can be used  
CC in diagnosing, preventing or treating diseases or conditions associated  
CC with the decreased expression or overexpression of the enzymes, such as  
CC autoimmune/inflammatory disorders (AIDS), allergies, atopic dermatitis,  
CC arthritis or thyroiditis), infections (e.g. bacterial, viral, parasitic,  
CC protozoal or fungal), metabolic disorders (e.g. obesity), reproductive  
CC disorders (e.g. infertility), neurological disorders (Parkinson's  
CC disease, Alzheimer's disease, stroke, epilepsy), cardiovascular disorders  
CC (e.g. myocardial infarction or hypertension), eye disorders, or cell  
CC proliferative diseases (e.g. cancer, atherosclerosis, hepatitis). They  
CC are also useful in assessing the effects of exogenous compounds on the  
CC expression of nucleic acid and amino acid sequences of the enzymes.  
XX  
XX  
XX Sequence 514 AA:

XX  
XX  
XX Query Match 72.7%; Score 32; DB 7; Length 514;  
XX Best Local Similarity 85.7%; Pred. No. 5.6e+02;  
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
XX 1 FLOWEPD 7  
XX |||||  
XX 307 FLOWEPD 313

RESULT 17  
 ADC24201  
 ID ADC24201 standard; protein, 551 AA.  
 XX  
 AC ADC24201;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Human NOV2b protein SEQ ID NO:8.  
 XX  
 KW human; NOV; cardiact; antiarteriosclerotic; hypotensive; vasotropic;  
 KW dermatological; anorectic; immunosuppressive; cytostatic;  
 KW anti-infectivity; haemostatic; anti-HIV; antiaesthetic; anti-inflammatory;  
 KW neuroprotective; anabolic; nootropic; antiparkinsonian; gene therapy;  
 KW cardiomyopathy; atherosclerosis; hypertension; congenital heart defect;  
 KW pulmonary stenosis; scleroderma; obesity; metabolic disturbance; obesity;  
 KW transplantaton; adrenoleukodystrophy; congenital adrenal hyperplasia;  
 KW prostate cancer; diabetes; metabolic disorder; neoplasm; adenocarcinoma;  
 KW fertility; haemophilia; graft versus host disease; AIDS;  
 KW bronchial asthma; Crohn's disease; multiple sclerosis;  
 KW infectious disease; anorexia; neurodegenerative disorder;  
 KW Alzheimer's disease; Parkinson's disease; immune disorder;  
 KW haematopoietic disorder; dyslipidaemia; wasting disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO2003076584-A2.  
 XX  
 PD 18-SEP-2003.  
 XX  
 PF 06-MAR-2003; 2003WO-US006951.  
 XX  
 PR 06-MAR-2002; 2002US-0361974P.  
 XX  
 PR 19-MAR-2002; 2002US-0365477P.  
 PR 22-MAR-2002; 2002US-036928P.  
 PR 06-AUG-2002; 2002US-0401661P.  
 PR 05-MAR-2003; 2003US-00401661.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Alsobrook JP, Burgess CE, Edinger SR, Gerlach VU, Ji W, Kekuda R;  
 PI Li L, Macdougall JR, Miller CE, Miller I, Paturajan M, Pena CBA;  
 PI Rieger DK, Sciore P, Shenoy SG, Smitson G, Spytek KA, Stone DJ;  
 PI Voss EZ, Zhong M;  
 XX  
 DR WPI; 2003-722330/68.  
 DR N-PSDB; ADC24200.  
 XX  
 PT New NOVX polypeptides and nucleic acids, useful for diagnosing or  
 PT treating e.g. cardiomyopathy, atherosclerosis, hypertension, scleroderma,  
 PT obesity, prostate cancer, AIDS, bronchial asthma, Crohn's disease, or  
 PT multiple sclerosis.  
 XX  
 PS Claim 1; SEQ ID NO 8; 229pp; English.

CC bronchial asthma, Crohn's disease, multiple sclerosis, infectious  
 CC disease, anorexia, neurodegenerative disorders (e.g. Alzheimer's disease,  
 CC or Parkinson's disease), immune disorders, haematopoietic disorders,  
 CC dyslipidaemias, and wasting disorders associated with chronic diseases.  
 CC The proteins can also be used as immunogens to produce antibodies and as  
 CC vaccines. The sequences may further be used in chromosome mapping,  
 CC identifying individual from minute biological samples (tissue typing),  
 CC and in forensic identification of a biological sample. The present  
 CC sequence represents human NOV2b from the present invention.  
 XX  
 SQ Sequence 551 AA;  
 XX  
 Query Match 72.7%; Score 32; DB 7; Length 551;  
 Best Local Similarity 85.7%; Pred. No. 6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Gaps 0;  
 XX  
 Qy 1 FLOLEFD 7  
 |||||  
 Db 307 FLOLEFD 313  
 XX  
 RESULT 18  
 ADD12554  
 ID ADD12554 standard; protein, 551 AA.  
 XX  
 AC ADD12554;  
 XX  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE Human ENZM-14 protein SEQ ID NO:14.  
 XX  
 KW human; enzyme; ENZM; cytostatic; antiarteriosclerotic; anti-HIV;  
 KW anti-allergic; cerebroprotective; antiparkinsonian; anticonvulsant;  
 KW nootropic; neuroprotective; anti-inflammatory; ophthalmological;  
 KW antihypertoid; antiarthritic; antibacterial; virucide; protozoacide;  
 KW antiparasitic; fungicide; anorectic; cardiact; hypotensive;  
 KW anti-infectivity; hepatotropic; gene therapy; autoimmune disorder;  
 KW inflammatory disorder; AIDS; allergy; atopic dermatitis; arthritis;  
 KW thyroiditis; infection; metabolic disorder; obesity;  
 KW reproductive disorder; infertility; neurological disorder;  
 KW Parkinson's disease; Alzheimer's disease; stroke; epilepsy;  
 KW cardiovascular disorder; myocardial infarction; hypertension;  
 KW eye disorder; cell proliferative disease; cancer; atherosclerosis;  
 KW hepatitis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003072729-A2.  
 XX  
 PD 04-SEP-2003.  
 XX  
 PF 21-FEB-2003; 2003WO-US005478.  
 XX  
 PR 22-FEB-2002; 2002US-0359513P.  
 PR 19-MAR-2002; 2002US-0365955P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Sprague WW, Kable AE, Lee S, Griffin JA, Becha SD, Lee EA;  
 PI Hafalia ADN, Khare R, Emerling BM, Marquis JP, Ramkumar J;  
 PI Elliott VS, Richardson TW, Baughn MR, Jin P, Chien D, Hawkins PR;  
 PI Wilson AD, Chawla NK, Tran UK, Lee SY, Zebaryjadian Y, Jiang X;  
 PI Jackson AA, Bhattacharya UG, Burdill JD, Blake JJ, Ho A, Zheng W;  
 XX  
 DR WPI; 2003-731608/69.  
 DR N-PSDB; ADD12611.  
 XX  
 PT New human enzymes and polynucleotides, useful for diagnosing, preventing  
 PT or treating diseases or conditions associated with aberrant enzyme  
 PT expression, e.g. cancer, hepatitis, AIDS, atherosclerosis, obesity,  
 PT infection or stroke.  
 XX  
 PS Claim 1; SEQ ID NO 14; 426pp; English.

XX The present sequence represents a human enzyme designated ENZM-14. Human  
 CC enzymes of the present invention have cytosolic, antiarteriosclerotic,  
 CC anti-HIV, antiallergic, cerebroprotective, antiparkinsonian,  
 CC anticonvulsant, nootropic, neuroprotective, antiinflammatory,  
 CC ophthalmological, antithyroid, antiarthritic, antibacterial, virucide,  
 CC protozoicidal, antiparasitic, fungicide, anorectic, cardiant, hypotensive,  
 CC antileukemic and hepatotropic activities, and can be used in gene  
 CC therapy. The human enzymes and polynucleotides encoding them can be used  
 CC in diagnosing, preventing or treating diseases or conditions associated  
 CC with the decreased expression or overexpression of the enzymes, such as  
 CC autoimmune/inflammatory disorders (AIDS, allergies, atopic dermatitis,  
 CC arthritis or thyroiditis), infections (e.g. bacterial, viral, parasitic,  
 CC protozoal or fungal), metabolic disorders (e.g. obesity), reproductive  
 CC disorders (e.g. infertility), neurological disorders (Parkinson's  
 CC disease, Alzheimer's disease, stroke, epilepsy), cardiovascular disorders  
 CC (e.g. myocardial infarction or hypertension), eye disorders, or cell  
 CC proliferative diseases (e.g. cancer, atherosclerosis, hepatitis). They  
 CC are also useful in assessing the effects of exogenous compounds on the  
 CC expression of nucleic acid and amino acid sequences of the enzymes.

Sequence 551 AA;

Query Match 72.7%; Score 32; DB 7; Length 551;  
 Best Local Similarity 85.7%; Pred. No. 6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEPD 7  
 |||||  
 Db 307 FLOLEPD 313

RESULT 19  
 AAY29648  
 ID AAY29648 standard; protein; 552 AA.

XX AAY29648;

DT 19-OCT-1999 (first entry)

DE Sheep prostaglandin H1 synthase query protein.

XX H-ras; HIV-1; reverse transcriptase; prostaglandin H2 synthase; PGHS-2;  
 KM protein folding inhibitor; PFI; growth; proliferation; drug.

XX Ovis sp.

XX Key Location/Qualifiers

FT Misc-difference 143 /note= "unspecified"  
 FT Misc-difference 144 /note= "unspecified"  
 FT Misc-difference 145 /note= "unspecified"  
 FT Misc-difference 146 /note= "unspecified"  
 FT Misc-difference 147 /note= "unspecified"  
 FT Misc-difference 148 /note= "unspecified"  
 FT Misc-difference 149 /note= "unspecified"  
 FT Misc-difference 150 /note= "unspecified"  
 FT Misc-difference 151 /note= "unspecified"  
 FT Misc-difference 152 /note= "unspecified"  
 FT Misc-difference 153 /note= "unspecified"  
 FT Misc-difference 154 /note= "unspecified"  
 FT Misc-difference 155 /note= "unspecified"

FT /note= "unspecified"  
 FT Misc-difference 166 /note= "unspecified"  
 FT Misc-difference 167 /note= "unspecified"  
 FT Misc-difference 168 /note= "unspecified"  
 FT Misc-difference 169 /note= "unspecified"  
 FT Misc-difference 170 /note= "unspecified"  
 FT Misc-difference 171 /note= "unspecified"  
 FT Misc-difference 172 /note= "unspecified"  
 FT Misc-difference 173 /note= "unspecified"  
 FT Misc-difference 174 /note= "unspecified"  
 FT Misc-difference 175 /note= "unspecified"  
 FT Misc-difference 176 /note= "unspecified"  
 FT Misc-difference 177 /note= "unspecified"  
 FT Misc-difference 178 /note= "unspecified"  
 FT Misc-difference 179 /note= "unspecified"

XX WO940435-A1.

XX 12-AUG-1999.

XX 06-FEB-1999; 99WO-US002612.

XX 09-FEB-1998; 98US-0074070P.

XX (NETZ/) NETZER W J.

XX Netzer WJ;

XX WPI; 1999-508524/42.

PT Methods for discovery, development and use of drugs and drug lead  
 PT molecules that inhibit protein folding.

PS Example; Page 37-38, 116pp; English.

XX The present invention describes the identification of protein folding  
 CC inhibitors (PFI) by determining the ratio of unfolded to folded protein  
 CC in the presence and absence of a test compound under protein biosynthetic  
 CC conditions after stress. Identifying PFI comprises: (1) contacting a  
 CC protein biosynthetic system under protein synthesis conditions with at  
 CC least one test compound; and (3) determining whether the test compound  
 CC increases the ratio of unfolded protein to folded protein, where an  
 CC increase in the ratio is indicative that the test compound is a PFI. The  
 CC method can be used for the discovery, development and use of drugs and  
 CC drug lead molecules that inhibit protein folding. The inhibitors  
 CC identified can be used in methods to inhibit cellular action of a protein  
 CC by inhibiting de novo folding in vivo, and optionally in conjunction with  
 CC heat shock treatment. The inhibitors can be used to modulate cellular  
 CC processes, e.g. enhancing the immunogenicity of a peptide or protein. The  
 CC inhibitors can modulate growth or proliferation of a cell by inhibiting  
 CC irreversible folding of the protein target. The tests can be carried out  
 CC using a wide range of procedures therefore making the detection more  
 CC simple and flexible for the operator. The present sequence represents  
 CC sheep prostaglandin H1 synthase (PGHS-1) used in the exemplification of  
 XX the present invention

Sequence 552 AA;

Query Match 72.7%; Score 32; DB 2; Length 552;

Best Local Similarity 85.7%; Pred. No. 6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 FLOLQEPD 7  
|||:|  
Db 323 FLOLQEPD 329

RESULT 20  
ADD12555  
ID ADD12555 standard; protein, 552 AA.  
AC ADD12555;  
XX 01-JAN-2004 (first entry)  
XX Human ENZM-15 protein SEQ ID NO:15.  
XX  
XX human; enzyme; ENZM; cytosolic; antiarteriosclerotic; anti-HIV;  
KW anti-atherogenic; cerebroprotective; antiparkinsonian; anticonvulsant;  
KW neurotrophic; neuroprotective; antiinflammatory; ophthalmological;  
KW antihypertensive; antidiabetic; antiparasitic; antitumor; antiproliferative;  
KW antiparasitic; fungicide; anorectic; cardiant; hypotensive;  
KW antileukemic; hepatocarcinoma; gene therapy; autoimmune disorder;  
KW inflammatory disorder; AIDS; allergy; atopic dermatitis; arthritis;  
KW thyroiditis; infection; metabolic disorder; obesity;  
KW reproductive disorder; infertility; neurological disorder;  
KW Parkinson's disease; Alzheimer's disease; stroke; epilepsy;  
KW cardiovascular disorder; myocardial infarction; hypertension;  
KW eye disorder; cell proliferative disease; cancer; atherosclerosis;  
KW hepatitis.  
XX Homo sapiens.  
XX MO2003072729-A2.  
XX 04-SEP-2003.  
XX 21-FEB-2003; 2003WO-US005478.  
XX 22-FEB-2002; 2002US-0359513P.  
XX 19-MAR-2002; 2002US-0365795P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Sprague WW, Kable AE, Lee S, Griffin JA, Becha SD, Lee EA;  
PI Hafalia AJA, Khare R, Emerling BM, Marquis JP, Rankumar J;  
PI Elliott VS, Richardson TW, Baughn MR, Jin P, Chien D, Hawkins PR;  
PI Wilson AD, Chawla NK, Tan UK, Lee SY, Zebardjian Y, Jiang X;  
PI Jackson AA, Bhatia UG, Burrill JD, Blake JJ, Ho A, Zheng W;  
XX WPI; 2003-731608/69.  
XX N-PSDB; ADD12612.  
XX  
XX New human enzymes and polynucleotides, useful for diagnosing, preventing  
XX or treating diseases or conditions associated with aberrant enzyme  
XX expression, e.g. cancer, hepatitis, AIDS, atherosclerosis, obesity,  
XX infection or stroke.  
XX  
XX Claim 1; SEQ ID NO 15; 426pp; English.  
XX  
XX The present sequence represents a human enzyme designated ENZM-15. Human  
XX enzymes of the present invention have cytosolic, antiarteriosclerotic,  
XX anti-HIV, anti-atherogenic, cerebroprotective, antiparkinsonian,  
XX anticonvulsant, neurotrophic, neuroprotective, antiinflammatory,  
XX ophthalmological, antidiabetic, antiparasitic, antitumor, antiproliferative,  
XX antiparasitic, fungicide, anorectic, cardiant, hypotensive,  
XX antileukemic, hepatocarcinoma, gene therapy, autoimmune disorder,  
XX inflammatory disorder, AIDS, allergy, atopic dermatitis, arthritis,  
XX thyroiditis, infection, metabolic disorder, obesity,  
XX reproductive disorder, infertility, neurological disorder,  
XX Parkinson's disease, Alzheimer's disease, stroke, epilepsy,  
XX cardiovascular disorder, myocardial infarction, hypertension,  
XX eye disorder, cell proliferative disease, cancer, atherosclerosis,  
XX hepatitis or thyroiditis, infections (e.g. bacterial, viral, parasitic,  
XX

CC protozoal or fungal), metabolic disorders (e.g. obesity), reproductive  
CC disorders (e.g. infertility), neurological disorders (Parkinson's  
CC disease, Alzheimer's disease, stroke, epilepsy), cardiovascular disorders  
CC (e.g. myocardial infarction or hypertension), eye disorders, or cell  
CC proliferative diseases (e.g. cancer, atherosclerosis, hepatitis). They  
CC are also useful in assessing the effects of exogenous compounds on the  
CC expression of nucleic acid and amino acid sequences of the enzymes.  
XX  
XX SQ Sequence 552 AA;

Query Match 72.7%; Score 32; DB 7; Length 552;  
Best Local Similarity 85.7%; Pred. No. 6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 FLOLQEPD 7  
|||:|  
Db 345 FLOLQEPD 351

RESULT 21  
ABB09701  
ID ABB09701 standard; protein, 562 AA.  
AC ABB09701;  
XX 11-JUN-2002 (first entry)  
XX  
XX Amino acid sequence of human COX-1b polypeptide.  
XX COX-1b; prostaglandin synthase-1; COX-1; prostaglandin; D2;  
KW 15-deoxy-Delta(12-14) J2; arachidonic acid; inflammation;  
KW rheumatoid arthritis; vascular disorder; skin disorder;  
KW venous insufficiency; atherosclerosis; skin ulcer; cancer;  
KW Alzheimer's disease; pain; migraine; cyclooxygenase; COX-2.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX Modified-site 67  
XX FT /note= "N-glycosylation site"  
XX Modified-site 103  
XX FT /note= "N-glycosylation site"  
XX Modified-site 143  
XX FT /note= "N-glycosylation site"  
XX  
XX FR2811677-A1.  
XX 18-JAN-2002.  
XX 12-JUL-2000; 2000FR-00009139.  
XX 12-JUL-2000; 2000FR-00009139.  
XX 12-JUL-2000; 2000FR-00009139.  
XX (INNO-) LAB INNOTHERA SA.  
XX Anger LM, Hanf R;  
XX WPI; 2002-149748/20.  
XX N-PSDB; ABL41898.  
XX  
XX Cell line that expresses isoform of prostaglandin synthase-1, useful for  
XX treating e.g. inflammation, converts arachidonic acid to antiinflammatory  
XX prostaglandins.  
XX  
XX Claim 4; Fig 6; 63pp; French.  
XX  
XX The present sequence represents a human COX-1b polypeptide. COX-1b  
XX represents an alternative splicing of the 3'-region of exon 9 of the  
XX mammalian gene for prostaglandin synthase-1 (COX-1). COX-1b  
XX preferentially produces the antiinflammatory prostaglandins D2 or 15-  
XX deoxy-Delta(12-14) J2 from arachidonic acid. COX-1b is used to treat  
XX inflammation (particularly rheumatoid arthritis); vascular and skin  
XX disorders (e.g. venous insufficiency, atherosclerosis or skin ulcers);

CC cancer (especially of the colon); Alzheimer's disease; pain and/or  
 CC migraine. Cells that express COX-1b are useful for screening compounds  
 CC for inhibition of cyclooxygenase activity, particularly for selective  
 CC action against COX-1, COX-1b or COX-2. Inhibitors are potentially useful  
 CC for treating the specified conditions

XX Sequence 562 AA;

Query Match  
 Best Local Similarity 72.7%; Score 32; DB 5; Length 562;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPD 7  
 ||||:|  
 DB 355 FLOLKFED 361

RESULT 22

ADD27944  
 ID ADD27944 standard; protein; 580 AA.

AC ADD27944;

DT 15-JAN-2004 (first entry)

DE COX-1 amino acid sequence SEQ ID NO:34.

XX cyclooxygenase type 1; cyclooxygenase type 1 variant protein;

KM COX-1 variant protein; genetic disease; tissue typing;

XX forensic identification.

XX Unidentified.

XX WO2003029411-A2.

XX 10-APR-2003.

PF 28-SEP-2002; 2002WO-US030947.

XX 28-SEP-2001; 2001US-0326133P.

PR 15-APR-2002; 2002US-0373225P.

PR 16-APR-2002; 2002US-0373661P.

PR 16-SEP-2002; 2002US-0411575P.

XX (UYYO ) UNIV BRIGHAM YOUNG.

PI Simmons D, Chandrasekharan VN;

DR WPI; 2003-421222/39.

XX Novel cyclooxygenase type 1 variant polypeptide and a polynucleotide  
 PT encoding the polypeptide, useful for identifying a compound that binds to  
 PT and modulates the activity of COX-1 variant polypeptide.

XX Disclosure; SEQ ID NO 34; 150p; English.

XX The present invention describes an isolated cyclooxygenase type 1 (COX-1)  
 CC variant polypeptide (I). (I) is useful for identifying a compound which  
 CC modulates the activity of (I). A nucleotide sequence encoding (I) can be  
 CC used for mapping their respective genes on a chromosome, and so locating  
 CC gene regions associated with genetic disease, identifying an individual  
 CC from a minute biological sample (tissue typing), and to aid in forensic  
 CC identification of a biological sample. The present sequence represents a  
 CC sequence which is used in the exemplification of the present invention.

XX Sequence 580 AA;

Query Match  
 Best Local Similarity 72.7%; Score 32; DB 7; Length 580;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPD 7  
 ||||:|

DB 344 FLOLKFED 350

RESULT 23  
 ADD27968  
 ID ADD27968 standard; protein; 580 AA.

AC ADD27968;

DT 15-JAN-2004 (first entry)

DE COX-1 amino acid sequence.

XX cyclooxygenase type 1; cyclooxygenase type 1 variant protein;

KM COX-1 variant protein; genetic disease; tissue typing;

XX forensic identification.

XX Unidentified.

XX WO2003029411-A2.

XX 10-APR-2003.

PF 28-SEP-2002; 2002WO-US030947.

XX 28-SEP-2001; 2001US-0326133P.

PR 15-APR-2002; 2002US-0373225P.

PR 16-APR-2002; 2002US-0373661P.

PR 16-SEP-2002; 2002US-0411575P.

XX (UYYO ) UNIV BRIGHAM YOUNG.

PI Simmons D, Chandrasekharan VN;

DR WPI; 2003-421222/39.

XX Novel cyclooxygenase type 1 variant polypeptide and a polynucleotide  
 PT encoding the polypeptide, useful for identifying a compound that binds to  
 PT and modulates the activity of COX-1 variant polypeptide.

XX Disclosure; Fig 2; 150p; English.

XX The present invention describes an isolated cyclooxygenase type 1 (COX-1)  
 CC variant polypeptide (I). (I) is useful for identifying a compound which  
 CC modulates the activity of (I). A nucleotide sequence encoding (I) can be  
 CC used for mapping their respective genes on a chromosome, and so locating  
 CC gene regions associated with genetic disease, identifying an individual  
 CC from a minute biological sample (tissue typing), and to aid in forensic  
 CC identification of a biological sample. The present sequence represents a  
 CC sequence which is used in the exemplification of the present invention.

XX Sequence 580 AA;

Query Match  
 Best Local Similarity 72.7%; Score 32; DB 7; Length 580;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPD 7  
 ||||:|  
 DB 344 FLOLKFED 350

RESULT 24

AA21690  
 ID AA21690 standard; protein; 599 AA.

AC AA21690;

DT 25-MAR-2003 (revised)

DT 23-JUL-1992 (first entry)

DE Prostaglandin endoperoxide synthase.

KW Human; hPBS; screening; anti-inflammatory; antibody.  
 XX  
 OS Homo sapiens.  
 XX JP04045786-A.  
 PN  
 XX 14-FEB-1992.  
 PD  
 XX 13-JUN-1990; 90JP-00152784.  
 PF  
 XX 13-JUN-1990; 90JP-00152784.  
 PR 13-JUN-1990; 90JP-00152784.  
 XX  
 PA (MEIP ) MEIUI MILK PROD CO LTD.  
 XX  
 DR WPI; 1992-101937/13.  
 DR N-PSDB; AAQ23001.  
 XX  
 PT Human prostaglandin endo-peroxide synthase - for screening  
 PT antiinflammatory agents.  
 XX  
 PS Claim 1; Page 1; 10pp; Japanese.  
 XX  
 CC The protein sequence of hPBS was deduced from the cDNA sequence obtd. by  
 CC screening a human genomic library in EMBL3. hPBS can be used to screen  
 CC anti-inflammatory agents. An anti-body against a peptide specific to hPBS  
 CC can be made, and used for the determin- ation of the protein. (Updated on  
 CC 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 599 AA;  
 SO  
 Query Match 72.7%; Score 32; DB 2; Length 599;  
 Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLOLEFD 7  
 DB 355 FLOLEKFD 361  
 BR  
 RESULT 25  
 ABG96403  
 ID ABG96403 standard; protein; 599 AA.  
 XX  
 AC ABG96403;  
 XX  
 DT 11-DEC-2002 (first entry)  
 XX  
 DE Human ovarian cancer marker OVS1.  
 XX  
 KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;  
 KW central nervous system disorder; bacterial meningitis; viral meningitis;  
 KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;  
 KW brain herniation; inflammation; encephalitis; testicular disorder;  
 KW nontuberculous granulomatous orchitis; connective tissue disorder;  
 KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;  
 KW histological type; carcinogenic; ovarian cancer marker.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200271928-A2.  
 XX  
 PD 19-SEP-2002.  
 XX  
 PF 14-MAR-2002; 2002WO-US007826.  
 XX  
 PR 14-MAR-2001; 2001US-0276025P.  
 PR 14-MAR-2001; 2001US-0276025P.  
 PR 10-AUG-2001; 2001US-0311732P.  
 PR 19-SEP-2001; 2001US-0323580P.  
 PR 26-SEP-2001; 2001US-0324967P.  
 PR 26-SEP-2001; 2001US-0325102P.  
 PR 26-SEP-2001; 2001US-0325149P.  
 XX

PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 XX Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;  
 PI Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;  
 PI Baer RC, Lu K, Schmandt RE, Zhao X, Glatt K;  
 XX  
 DR WPI; 2002-723277/78.  
 DR N-PSDB; ABS76502.  
 XX  
 PT Assessing whether a patient is afflicted with ovarian cancer, useful in  
 PT assessing the stage or progression of the disease, comprises comparing  
 PT the expression level of a cancer marker in a sample from a patient and  
 PT from a non cancer patient.  
 XX  
 PS Disclosure; Page 386-388; 481pp; English.  
 XX  
 CC The present invention relates to a new method for assessing whether a  
 CC patient is afflicted with ovarian cancer. The method involves comparing  
 CC the expression level of a marker in a patient sample and the normal level  
 CC of expression of the marker in a control non-ovarian cancer sample, where  
 CC the marker is selected from 363 cancer markers described in the  
 CC specification. The method of the invention is useful in diagnosing or  
 CC characterising cancer, in detecting the presence of cancer as early as  
 CC possible, and the recurrence of ovarian cancer. The method may also be of  
 CC particular use with patients having an enhanced risk of developing  
 CC ovarian cancer (e.g. patients having a familial history of ovarian  
 CC cancer). The cancer markers may be used in the management and treatment  
 CC of e.g. brain and central nervous system disorders (e.g. bacterial and  
 CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain  
 CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),  
 CC inflammations (e.g. bacterial or viral meningitis or encephalitis),  
 CC testicular disorders (e.g. nontuberculous granulomatous orchitis),  
 CC connective tissue disorders, or heart disorders (e.g. ischaemic heart  
 CC disease or atherosclerosis). The compositions and methods may also be  
 CC used in assessing the histological type of neoplasm associated with  
 CC ovarian cancer, monitoring the progression of ovarian cancer, determining  
 CC whether ovarian cancer has metastasized or is likely to metastasize,  
 CC selecting a composition for inhibiting ovarian cancer, assessing the  
 CC ovarian carcinogenic potential of a compound, or inhibiting ovarian  
 CC cancer or at risk of developing ovarian cancer. The present amino acid  
 CC sequence represents one of the ovarian cancer markers described in the  
 CC invention.  
 XX  
 SQ Sequence 599 AA;  
 SO  
 Query Match 72.7%; Score 32; DB 5; Length 599;  
 Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLOLEFD 7  
 DB 355 FLOLEKFD 361  
 BR  
 RESULT 26  
 ABR42249  
 ID ABR42249 standard; protein; 599 AA.  
 XX  
 AC ABR42249;  
 XX  
 DT 28-JUL-2003 (first entry)  
 XX  
 DE Human cyclooxygenase 1 (COX-1).  
 XX  
 KW Cyclooxygenase 1; COX-1; human; enzyme; vulnerary; osteopathic;  
 KW gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003022224-A2.  
 XX  
 PD 20-MAR-2003.  
 XX



Pf	11-SEP-2002; 2002WO-US028930.
Xx	
Pr	11-SEP-2001; 2001US-00953067.
Xx	
Pa	(UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
Xx	
Pi	O'Connor PJ;
Xx	
Dr	WPI, 2003-313185/30.
Dr	N-PSDB; ACC57773.
Pt	Novel vector useful for enhancing wound healing or treating osteoporosis,
Pt	osteogenesis imperfecta, and brittle bone conditions, comprises a
Pt	promoter linked to a cyclooxygenase expression cassette.
Xx	
Ps	Disclosure; Page 69-72; 88pp; English.
Xx	
Cc	The present sequence is the protein sequence of human cyclooxygenase 1
Cc	(COX-1). A claimed vector for use in enhancing wound healing comprises a
Cc	promoter linked to a COX expression cassette. A claimed method for
Cc	enhanced wound healing involves delivering the vector to the location of
Cc	the wound. The wound may be a bone fracture or a skin wound. A claimed
Cc	method for treating osteoporosis, osteogenesis imperfecta and brittle
Cc	bone conditions comprises administering the vector. A claimed composition
Cc	for use in wound healing comprises a COX protein, especially COX-1, COX-2
Cc	or both
Sq	
Xx	Sequence 599 AA;
Qy	
Dz	
Db	Query Match 72.7%; Score 32; DB 6; Length 599; Best Local Similarity 85.7%; Pred. No. 6.6e+02; Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0.  1 FLDLEPD 7    :   355 FLQLKFD 361
Result 27	
ADCC24199	
ID ADC24199	standard; protein; 599 AA.
XX	
AC	ADC24199;
DT	
DT	18-DEC-2003 (first entry)
DE	
Human NOV2a	protein SEQ ID NO:6.
Kw	human; NOX; cardiac; antiarteriosclerotic; hypotensive; vasotropic;
Kw	dematological; anorectic; immunosuppressive; cytostatic;
Kw	antileukemia; haemostatic; anti-HIV; antiaesthetic; antiinflammatory;
Kw	neuroprotective; anabolic; nootropic; antiparkinsonian; gene therapy;
Kw	cardiomyopathy; atherosclerosis; hypertension; congenital heart defect;
Kw	pulmonary stenosis; scleroderma; obesity; metabolic disturbance; obesity;
Kw	transplantation; adrenoleukodystrophy; congenital adrenal hyperplasia;
Kw	prostate cancer; diabetes; metabolic disorder; neoplasm; adenocarcinoma;
Kw	fertility; haemophilia; graft versus host disease; AIDS;
Kw	bronchial asthma; Crohn's disease; multiple sclerosis;
Kw	infectious disease; anorexia; neurodegenerative disorder;
Kw	Alzheimer's disease; Parkinson's disease; immune disorder;
Kw	haematopoietic disorder; dyslipidaemia; wasting disorder.
XX	
Os	Homo sapiens.
XX	
Pn	WO2003076584-A2.
Xx	
Pd	18-SEP-2003.
Xx	
Pf	06-MAR-2003; 2003WO-US006951.
Pr	06-MAR-2002; 2002US-0361974P.
Pr	19-MAR-2002; 2002US-0365477P.
Pr	22-MAR-2002; 2002US-0366928P.

XX	PR	06-AUG-2002; 2002US-0401661P.
XX	PR	05-MAR-2003; 2003US-00401661.
XX	PA	(CURA-) CURAGEN CORP.
XX	PI	Alabrook JP, Burgess CE, Edinger SR, Gerlach VL, Ji W, Kekuda R;
XX	PI	Li L, Macdonnell JR, Miller CE, Millet I, Paturjan M, Pena CEA;
XX	PI	Rieger DK, Sciore P, Shenoy SG, Smithson G, Spytek KA, Stone DJ,
XX	PI	Voss BZ, Zhong M;
XX	XX	WPI; 2003-722330/68.
DR	N-PSDB	ADDC24198.
XX	FT	New NOVX polypeptides and nucleic acids useful for diagnosing or
XX	FT	treating e.g. cardiomyopathy, atherosclerosis, hypertension, scleroderma,
XX	FT	obesity, prostate cancer, AIDS, bronchial asthma, Crohn's disease, or
XX	FT	multiple sclerosis.
XX	PS	Claim 1; SEQ ID NO 6; 22pp; English.
XX	CC	The present invention describes novel human proteins, designated NOVX
XX	CC	proteins. The NOVX sequences have cardiant, antiatherosclerotic,
XX	CC	hypotensive, vasotropic, dermatologic, anorectic, immunosuppressive,
XX	CC	cycostatic, antifertility, haemostatic, anti-HIV, antislutamic,
XX	CC	antiinflammatory, neuroprotective, anabolic, nootropic and
XX	CC	antiparkinsonian activities, and can be used in gene therapy. The NOVX
XX	CC	sequences can be used as a therapeutic in the manufacture of a medicament
XX	CC	for treating a syndrome associated with a human disease, such as a
XX	CC	pathology associated with NOVX. The NOVX proteins and nucleic acids
XX	CC	encoding them are useful for diagnosing or treating pathologies, diseases
XX	CC	or conditions associated with NOVX sequences, including cardiomyopathy,
XX	CC	atherosclerosis, hypertension, congenital heart defects, pulmonary
XX	CC	stenosis, scleroderma, obesity, metabolic disturbances associated with
XX	CC	obesity, transpiration, adrenoleukodystrophy, congenital adrenal
XX	CC	hyperplasia, prostate cancer, diabetes, metabolic disorders, neoplasm,
XX	CC	adenocarcinoma, fertility, haemophilia, graft versus host disease, AIDS,
XX	CC	bronchial asthma, Crohn's disease, multiple sclerosis, infectious
XX	CC	disease, anorexia, neurodegenerative disorders (e.g. Alzheimer's disease,
XX	CC	or Parkinson's disease), immune disorders, haematopoietic disorder,
XX	CC	dyliademiatis, and waiting disorders associated with chronic diseases.
XX	CC	The proteins can also be used as immunogens to produce antibodies and as
XX	CC	vaccines. The sequences may further be used in chromosome mapping,
XX	CC	identifying individual from minute biological samples (tissue typing),
XX	CC	and in forensic identification of a biological sample. The present
XX	CC	sequence represents human NOV2a from the present invention.
XX	SQ	Sequence 599 AA:
XX	Query Match	72.7%; Score 32; DB 7; Length 599;
XX	Best Local Similarity	85.7%; Pred. No. 6.6e+02;
XX	Matches	6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY	1 FLOWEPD 7	
bj		
	355 FLOWEPD 361	
RESULT 28		
ADD27936	ID	ADD27936 standard; protein; 599 AA.
AC	ADD27936;	
XX	DT	15-JAN-2004 (first entry)
DE	Human COX-1 amino acid sequence SEQ ID NO:22.	
XX	Cytochrome c oxidase type 1; cytochrome c oxidase type 1 variant protein;	
KM	COX-1 variant protein; genetic disease; tissue typing;	
XX	forensic identification.	
OS	Homo sapiens.	

PN WO2003029411-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 XX 28-SEP-2002; 2002WO-US030947.  
 PF  
 XX 28-SEP-2001; 2001US-0326133P.  
 PR 15-APR-2002; 2002US-0373225P.  
 PR 16-APR-2002; 2002US-0373661P.  
 PR 16-SEP-2002; 2002US-0411575P.  
 XX  
 PA (UYO ) UNIV BRIGHAM YOUNG.  
 XX  
 XX Simmons D, Chandrasekharan VN;  
 PI  
 DR WPI; 2003-421222/39.  
 XX  
 PT Novel cyclooxygenase type 1 variant polypeptide and a polynucleotide  
 PT encoding the polypeptide, useful for identifying a compound that binds to  
 PT and modulates the activity of COX-1 variant polypeptide.  
 XX  
 PS Disclosure; SEQ ID NO 22; 150pp; English.  
 XX  
 CC The present invention describes an isolated cyclooxygenase type 1 (COX-1)  
 CC variant polypeptide (I). (I) is useful for identifying a compound which  
 CC modulates the activity of (I). A nucleotide sequence encoding (I) can be  
 CC used for mapping their respective genes on a chromosome, and so locating  
 CC gene regions associated with genetic disease, identifying an individual  
 CC from a minute biological sample (tissue typing), and to aid in forensic  
 CC identification of a biological sample. The present sequence represents a  
 CC sequence which is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 599 AA;  
 XX  
 QY  
 Db 1 FLOLQKFD 7  
 355 FLOLQKFD 361  
 Query Match 72.7%; Score 32; DB 7; Length 599;  
 Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

PT New PRO nucleic acid or polypeptide, useful for preparing a  
 PT pharmaceutical composition for diagnosing or treating psoriasis in a  
 PT mammal.  
 PD  
 XX  
 XX Claim 9; SEQ ID NO 2020; 3069pp; English.  
 PS  
 CC The invention relates to novel polynucleotide and polypeptides for  
 CC treating psoriasis or a sequence having at least 80% identity to the  
 CC above sequences. The nucleic acid is useful for preparing a composition  
 CC for diagnosing or treating psoriasis in a mammal. This sequence  
 CC corresponds to one of the polypeptides of the invention.  
 XX  
 SQ Sequence 599 AA;  
 XX  
 QY  
 Db 1 FLOLQKFD 7  
 355 FLOLQKFD 361  
 Query Match 72.7%; Score 32; DB 8; Length 599;  
 Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 30  
 AAP91008  
 ID AAP91008 standard; protein; 600 AA.  
 XX  
 XX AAP91008;  
 AC  
 XX 25-MAR-2003 (revised)  
 DT 10-MAR-2003 (revised)  
 DT 13-MAR-1990 (first entry)  
 XX  
 XX Prostaglandin endoperoxide synthetase.  
 DE  
 XX  
 XX Recombinant.  
 KW  
 OS Ovis aries.  
 XX  
 XX JP01228479-A.  
 PN  
 XX  
 PD 12-SEP-1989.  
 XX  
 XX 10-MAR-1988; 88JP-00054941.  
 PF  
 XX 10-MAR-1988; 88JP-00054941.  
 PR  
 XX 10-MAR-1988; 88JP-00054941.  
 XX  
 PA (MEIP ) MEIJI MILK PROD CO LTD.  
 XX  
 XX WPI; 1989-312224/43.  
 DR N-PSDB; AAN91622.  
 DR  
 XX Prostaglandin endo-peroxide synthetase coding DNA - useful as starting  
 PT material for recombinant, and for synthesis of various prostaglandin(s)  
 PT and laboratory reagent.  
 PT  
 XX  
 XX Claim 1; Fig 1; 11pp; Japanese.  
 PS  
 CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003  
 CC to correct PA field.)  
 CC  
 XX Sequence 600 AA;  
 SQ  
 Query Match 72.7%; Score 32; DB 1; Length 600;  
 Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLQKFD 7  
 Db 356 FLOLQKFD 362  
 RESULT 31

ADD27937  
 ID ADD27937 standard; protein; 600 AA.  
 AC ADD27937;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Ovine COX-1 amino acid sequence SEQ ID NO:23.  
 XX  
 KM cyclooxygenase type 1; cyclooxygenase type 1 variant protein;  
 KM COX-1 variant protein; genetic disease; tissue typing;  
 KM forensic identification.  
 XX  
 OS Ovis sp.  
 XX  
 PN MO2003029411-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 PF 28-SEP-2002; 2002MO-US030947.  
 XX  
 PR 28-SEP-2001; 2001US-0326133P.  
 PR 15-APR-2002; 2002US-0373225P.  
 PR 16-APR-2002; 2002US-0373661P.  
 PR 16-SEP-2002; 2002US-0411575P.  
 PA (UYYO ) UNIV BRIGHAM YOUNG.  
 XX  
 PI Simmons D, Chandrasekharan VN;  
 XX  
 DR WPI; 2003-421222/39.  
 XX  
 PT Novel cyclooxygenase type 1 variant polypeptide and a polynucleotide  
 PT encoding the polypeptide, useful for identifying a compound that binds to  
 PT and modulates the activity of COX-1 variant polypeptide.  
 CC  
 CC Disclosure; SEQ ID NO 23; 150bp; English.  
 XX  
 SQ  
 Query Match 72.7%; Score 32; DB 7; Length 600;  
 Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLQLKFD 7  
 DB 356 FLQLKFD 362

RESULT 32  
 ABB57303  
 ID ABB57303 standard; protein; 602 AA.  
 XX  
 AC ABB57303;  
 XX  
 DT 07-MAR-2002 (first entry)  
 XX  
 DE Mouse ischaemic condition related protein sequence SEQ ID NO:849.  
 XX  
 KM Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
 KM vasospastic ischaemia; ischaemic condition; ischaemic disease.  
 XX  
 OS Mus musculus.  
 XX

PN MO200188188-A2.  
 XX  
 PD 22-NOV-2001.  
 XX  
 PF 18-MAY-2001; 2001MO-JP004192.  
 XX  
 PR 18-MAY-2000; 2000JP-00145977.  
 XX  
 PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.  
 XX  
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;  
 XX  
 DR WPI; 2002-034733/04.  
 DR N-PSDB; ABI99766.  
 XX  
 PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
 PT expression levels of particular genes defined in the specification or by  
 PT determining the expression profile of a gene group comprising these  
 PT genes.  
 XX  
 PS Claim 2; Page 2109-2111; 2690bp; English.  
 XX  
 CC The present invention describes a method for examining ischaemic  
 CC conditions, comprising measuring the expression levels of particular  
 CC genes (1) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (1). The method  
 CC is useful for examining the ischaemic condition (e.g. compressive  
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding the  
 CC protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The expression  
 CC levels or expression profiles produced by these genes are used as an  
 CC indicator when screening for ischaemic condition-improving drugs or  
 CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR  
 CC primers for a mouse ischaemic condition related sequence, which are used  
 CC in the exemplification of the present invention  
 XX  
 SQ  
 Query Match 72.7%; Score 32; DB 5; Length 602;  
 Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLQLKFD 7  
 DB 358 FLQLKFD 364

RESULT 33  
 ABR42251  
 ID ABR42251 standard; protein; 602 AA.  
 XX  
 AC ABR42251;  
 XX  
 DT 28-JUL-2003 (first entry)  
 XX  
 DE Mouse cyclooxygenase 1 (COX-1).  
 XX  
 KM Cyclooxygenase 1; COX-1; mouse; enzyme; vulnerrary; osteopathic;  
 KM gene therapy.  
 XX  
 OS Mus sp.  
 XX  
 PN MO2003022224-A2.  
 XX  
 PD 20-MAR-2003.  
 XX  
 PF 11-SEP-2002; 2002MO-US028930.  
 XX  
 PR 11-SEP-2001; 2001US-00953067.  
 XX  
 PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.  
 XX

PI O'Connor PJ;  
 XX WPI; 2003-333185/30.  
 DR N-PSDB; ACC57775.  
 XX  
 PT Novel vector useful for enhancing wound healing or treating osteoporosis,  
 PT osteogenesis imperfecta, and brittle bone conditions, comprises a  
 PT promoter linked to a cyclooxygenase expression cassette.  
 XX  
 PS Disclosure; Page 81-84; 88pp; English.  
 XX  
 CC The present sequence is the protein sequence of mouse cyclooxygenase 1  
 CC (COX-1). A claimed vector for use in enhancing wound healing comprises a  
 CC promoter linked to a COX expression cassette. A claimed method for  
 CC enhanced wound healing involves delivering the vector to the location of  
 CC the wound. The wound may be a bone fracture or a skin wound. A claimed  
 CC method for treating osteoporosis, osteogenesis imperfecta and brittle  
 CC bone conditions comprises administering the vector. A claimed composition  
 CC for use in wound healing comprises a COX protein, especially COX-1, COX-2  
 CC or both.  
 CC  
 SQ Sequence 602 AA;  
 XX  
 XX  
 Query Match 72.7%; Score 32; DB 6; Length 602;  
 Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLOLEFPD 7  
 |||||  
 Db 358 FLOLEKFD 364  
 XX  
 XX  
 RESULT 34  
 ADD27941  
 ID ADD27941 standard; protein; 602 AA.  
 XX  
 AC ADD27941;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Rat COX-1 amino acid sequence SEQ ID NO:30.  
 XX  
 KW Cyclooxygenase type 1; cyclooxygenase type 1 variant protein;  
 KW COX-1 variant protein; genetic disease; tissue typing;  
 KW forensic identification.  
 XX  
 OS Rattus sp.  
 XX  
 PN WO2003029411-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 PF 28-SEP-2002; 2002WO-US030947.  
 XX  
 PR 28-SEP-2001; 2001US-0326133P.  
 PR 15-APR-2002; 2002US-0373225P.  
 PR 16-APR-2002; 2002US-0373661P.  
 PR 16-SEP-2002; 2002US-0411575P.  
 XX  
 PA (UYYO) UNIV BRIGHAM YOUNG.  
 XX  
 PI Simmons D, Chandrasekharan VN;  
 XX  
 DR WPI; 2003-421222/39.  
 XX  
 PT Novel cyclooxygenase type 1 variant polypeptide and a polynucleotide  
 PT encoding the polypeptide, useful for identifying a compound that binds to  
 PT and modulates the activity of COX-1 variant polypeptide.  
 XX  
 PS Disclosure; SEQ ID NO 30; 150pp; English.  
 XX  
 CC The present invention describes an isolated cyclooxygenase type 1 (COX-1)  
 CC variant polypeptide (I). (I) is useful for identifying a compound which

CC modulates the activity of (I). A nucleotide sequence encoding (I) can be  
 CC used for mapping their respective genes on a chromosome, and so locating  
 CC gene regions associated with genetic disease, identifying an individual  
 CC from a minute biological sample (tissue typing), and to aid in forensic  
 CC identification of a biological sample. The present sequence represents a  
 CC sequence which is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 602 AA;  
 XX  
 XX  
 Query Match 72.7%; Score 32; DB 7; Length 602;  
 Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLOLEFPD 7  
 |||||  
 Db 358 FLOLEKFD 364  
 XX  
 XX  
 RESULT 35  
 ADD27940  
 ID ADD27940 standard; protein; 602 AA.  
 XX  
 AC ADD27940;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Murine COX-1 amino acid sequence SEQ ID NO:29.  
 XX  
 KW Cyclooxygenase type 1; cyclooxygenase type 1 variant protein;  
 KW COX-1 variant protein; genetic disease; tissue typing;  
 KW forensic identification.  
 XX  
 OS Mus sp.  
 XX  
 PN WO2003029411-A2.  
 XX  
 PD 10-APR-2003.  
 XX  
 PF 28-SEP-2002; 2002WO-US030947.  
 XX  
 PR 28-SEP-2001; 2001US-0326133P.  
 PR 15-APR-2002; 2002US-0373225P.  
 PR 16-APR-2002; 2002US-0373661P.  
 PR 16-SEP-2002; 2002US-0411575P.  
 XX  
 PA (UYYO) UNIV BRIGHAM YOUNG.  
 XX  
 PI Simmons D, Chandrasekharan VN;  
 XX  
 DR WPI; 2003-421222/39.  
 XX  
 PT Novel cyclooxygenase type 1 variant polypeptide and a polynucleotide  
 PT encoding the polypeptide, useful for identifying a compound that binds to  
 PT and modulates the activity of COX-1 variant polypeptide.  
 XX  
 PS Disclosure; SEQ ID NO 29; 150pp; English.  
 XX  
 CC The present invention describes an isolated cyclooxygenase type 1 (COX-1)  
 CC variant polypeptide (I). (I) is useful for identifying a compound which  
 CC modulates the activity of (I). A nucleotide sequence encoding (I) can be  
 CC used for mapping their respective genes on a chromosome, and so locating  
 CC gene regions associated with genetic disease, identifying an individual  
 CC from a minute biological sample (tissue typing), and to aid in forensic  
 CC identification of a biological sample. The present sequence represents a  
 CC sequence which is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 602 AA;  
 XX  
 XX  
 Query Match 72.7%; Score 32; DB 7; Length 602;  
 Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLOLEFPD 7

Db 358 FLQLKFD 364

RESULT 36  
ADD27938  
ADD27938 standard; protein; 603 AA.

AC ADD27938;  
DT 15-JAN-2004 (first entry)

DE Canine COX-1 amino acid sequence SEQ ID NO:24.

KW Cyclooxygenase type 1; cyclooxygenase type 1 variant protein;  
CX COX-1 variant protein; genetic disease; tissue typing;  
KM forensic identification.

XX Canis sp.

XX WO2003029411-A2.

XX 10-APR-2003.

XX 28-SEP-2002; 2002WO-US030947.

XX 28-SEP-2001; 2001US-0326133P.

XX 15-APR-2002; 2002US-0373225P.

XX 16-APR-2002; 2002US-0373661P.

XX 16-SEP-2002; 2002US-0411575P.

XX (UYUO) UNITV BRIGHAM YOUNG.

XX Simmons D, Chandrasekharan VN;

XX WPI; 2003-421222/39.

PT Novel cyclooxygenase type 1 variant polypeptide and a polynucleotide  
encoding the polypeptide, useful for identifying a compound that binds to  
and modulates the activity of COX-1 variant polypeptide.

PS Disclosure; SEQ ID NO 24; 150pp; English.

CC The present invention describes an isolated cyclooxygenase type 1 (COX-1)  
CC variant polypeptide (I). (I) is useful for identifying a compound which  
CC modulates the activity of (I). A nucleotide sequence encoding (I) can be  
CC used for mapping their respective genes on a chromosome, and so locating  
CC gene regions associated with genetic disease, identifying an individual  
CC from a minute biological sample (tissue typing), and to aid in forensic  
CC identification of a biological sample. The present sequence represents a  
CC sequence which is used in the exemplification of the present invention.

XX Sequence 603 AA;

Query Match 72.7%; Score 32; DB 7; Length 603;  
Best Local Similarity 85.7%; Pred. No. 6.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLQLKFD 7  
Db 359 FLQLKFD 365

RESULT 37  
ADD27939

ID ADD27939 standard; protein; 606 AA.

AC ADD27939;

DT 15-JAN-2004 (first entry)

DE Rabbit COX-1 amino acid sequence SEQ ID NO:27.

KW Cyclooxygenase type 1; cyclooxygenase type 1 variant protein;  
CX COX-1 variant protein; genetic disease; tissue typing;  
KM forensic identification.

XX Oryctolagus cuniculus.

XX WO2003029411-A2.

XX 10-APR-2003.

XX 28-SEP-2002; 2002WO-US030947.

XX 28-SEP-2001; 2001US-0326133P.

XX 15-APR-2002; 2002US-0373225P.

XX 16-APR-2002; 2002US-0373661P.

XX 16-SEP-2002; 2002US-0411575P.

XX (UYUO) UNITV BRIGHAM YOUNG.

XX Simmons D, Chandrasekharan VN;

XX WPI; 2003-421222/39.

PT Novel cyclooxygenase type 1 variant polypeptide and a polynucleotide  
encoding the polypeptide, useful for identifying a compound that binds to  
and modulates the activity of COX-1 variant polypeptide.

CC The present invention describes an isolated cyclooxygenase type 1 (COX-1)  
CC variant polypeptide (I). (I) is useful for identifying a compound which  
CC modulates the activity of (I). A nucleotide sequence encoding (I) can be  
CC used for mapping their respective genes on a chromosome, and so locating  
CC gene regions associated with genetic disease, identifying an individual  
CC from a minute biological sample (tissue typing), and to aid in forensic  
CC identification of a biological sample. The present sequence represents a  
CC sequence which is used in the exemplification of the present invention.

XX Sequence 606 AA;

Query Match 72.7%; Score 32; DB 7; Length 606;  
Best Local Similarity 85.7%; Pred. No. 6.7e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLQLKFD 7  
Db 362 FLQLKFD 368

RESULT 38  
ABB07241

ID ABB07241 standard; protein; 608 AA.

AC ABB07241;

DT 26-MAR-2002 (first entry)

DE Canine cyclooxygenase-1 (COX-1) protein.

KW Cyclooxygenase; COX-1; COX-2; canine; arthritis; cancer; neoplasia;  
CX inflammation; central nervous system; dog.

XX Canis familiaris.

XX Location/Qualifiers

FT Key 15..34  
FT MISC-difference 15..34  
FT /note="residues Xaa are unknown; the encoding codons are  
not indicated in the corresponding DNA sequence"

PN WO200111026-A1.

PD 15-FEB-2001.

PF 04-AUG-2000; 2000WO-US019565.  
XX  
PR 06-AUG-1999; 99US-0147601P.  
XX  
PA (SEAR ) SEARLE & CO G D.  
XX  
PI Gierse JK;  
XX  
DR WPI; 2002-113777/15.  
XX  
DR N-PSDB; ABA94341.  
XX  
PT Novel substantially purified canine cyclooxygenase 1 or 2 protein, useful  
PT for identifying drugs that can reduce inflammation in dogs, and screening  
PT selective inhibitors of cyclooxygenase-2 protein.  
XX  
PS Claim 11; Page 81-83; 122pp; English.  
XX  
CC The invention relates to genes that encode canine cyclooxygenase (COX)-1  
CC or COX-2 proteins. The COX proteins, especially COX-2 is useful for  
CC diagnosing or prognosing a COX-2 related condition in a dog. COX-2 or its  
CC fragment is useful for identifying a test material which has the ability  
CC to inhibit, suppress, modulate, or maintain canine COX-2 activity. The  
CC COX-1 and COX-2 polynucleotides are useful for determining an association  
CC between a polymorphism and a trait. COX-2 cDNA molecules and methods  
CC provided are also useful for diagnosing or prognosing COX-2 related  
CC condition such as arthritis, cancer, neoplasia, inflammation or central  
CC nervous system disorder in a dog. The present sequence represents a  
CC canine COX-1 protein  
XX  
SQ Sequence 608 AA;  
XX  
Query Match 72.7%; Score 32; DB 5; Length 608;  
Best Local Similarity 85.7%; Pred. No. 6.7e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 FLOLEFD 7  
DB 364 FLOLEFD 370  
XX  
RESULT 39  
AB07243 standard; protein; 608 AA.  
XX  
AC AB07243;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human cyclooxygenase-1 (COX-1) protein.  
XX  
KW Cyclooxygenase; COX-1; COX-2; canine; arthritis; cancer; neoplasia;  
KW inflammation; central nervous system; human.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 4..39  
FT /note= "residues Xaa are unknown; the encoding codons are  
FT not indicated in the corresponding DNA sequence"  
FT Misc-difference 105  
FT /note= "encoded by CTC"  
XX  
PN WO200111026-A1.  
XX  
PD 15-FEB-2001.  
XX  
PF 04-AUG-2000; 2000WO-US019565.  
XX  
PR 06-AUG-1999; 99US-0147601P.  
XX  
PS (SEAR ) SEARLE & CO G D.  
XX  
PI Gierse JK;  
XX

XX  
DR WPI; 2002-113777/15.  
XX  
DR N-PSDB; ABA94343.  
XX  
PT Novel substantially purified canine cyclooxygenase 1 or 2 protein, useful  
PT for identifying drugs that can reduce inflammation in dogs, and screening  
PT selective inhibitors of cyclooxygenase-2 protein.  
XX  
PS Disclosure; Page 92-94; 122pp; English.  
XX  
CC The invention relates to genes that encode canine cyclooxygenase (COX)-1  
CC or COX-2 proteins. The COX proteins, especially COX-2 is useful for  
CC diagnosing or prognosing a COX-2 related condition in a dog. COX-2 or its  
CC fragment is useful for identifying a test material which has the ability  
CC to inhibit, suppress, modulate, or maintain canine COX-2 activity. The  
CC COX-1 and COX-2 polynucleotides are useful for determining an association  
CC between a polymorphism and a trait. COX-2 cDNA molecules and methods  
CC provided are also useful for diagnosing or prognosing COX-2 related  
CC condition such as arthritis, cancer, neoplasia, inflammation or central  
CC nervous system disorder in a dog. The present sequence represents a human  
CC COX-1 protein  
XX  
SQ Sequence 608 AA;  
XX  
Query Match 72.7%; Score 32; DB 5; Length 608;  
Best Local Similarity 85.7%; Pred. No. 6.7e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 FLOLEFD 7  
DB 364 FLOLEFD 370  
XX  
RESULT 40  
ADD27929  
ID ADD27929 standard; protein; 629 AA.  
XX  
AC ADD27929;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Human hCOX-3(cs) reading frame 2 protein SEQ ID NO:15.  
XX  
KW cyclooxygenase type 1; cyclooxygenase type 1 variant protein;  
KW COX-1 variant protein; genetic disease; tissue typing;  
KW forensic identification; human; COX-3.  
XX  
OS Homo sapiens.  
XX  
PN WO2003029411-A2.  
XX  
PD 10-APR-2003.  
XX  
PF 28-SEP-2002; 2002WO-US030947.  
XX  
PR 28-SEP-2001; 2001US-0326133P.  
PR 15-APR-2002; 2002US-0373225P.  
PR 16-APR-2002; 2002US-0373661P.  
PR 16-SEP-2002; 2002US-0411575P.  
XX  
PA (UYIO ) UNIV BRIGHAM YOUNG.  
XX  
PI Simmons D, Chandrasekharan VN;  
XX  
DR WPI; 2003-421222/39.  
XX  
PT Novel cyclooxygenase type 1 variant polypeptide and a polynucleotide  
PT encoding the polypeptide, useful for identifying a compound that binds to  
PT and modulates the activity of COX-1 variant polypeptide.  
XX  
PS Claim 18; SEQ ID NO 15; 150pp; English.  
XX  
CC The present invention describes an isolated cyclooxygenase type 1 (COX-1)

CC variant polypeptide (I). (I) is useful for identifying a compound which  
 CC modulates the activity of (I). A nucleotide sequence encoding (I) can be  
 CC used for mapping their respective genes on a chromosome, and so locating  
 CC gene regions associated with genetic disease, identifying an individual  
 CC from a minute biological sample (tissue typing), and to aid in forensic  
 CC identification of a biological sample. The present sequence represents a  
 CC sequence which is used in the exemplification of the present invention.

XX  
 SQ Sequence 629 AA;

Query Match 72.7%; Score 32; DB 7; Length 629;  
 Best Local Similarity 85.7%; Pred. No 6.9e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQLKPD 7  
 |||: ||  
 Db 385 FLQLKPD 391

Search completed: January 12, 2005, 20:06:49  
 Job time : 204.6 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 20:16:00 ; Search time 93.6 Seconds  
(without alignments)  
34.739 Million cell updates/sec

Title: US-09-870-216C-5  
Perfect score: 44  
Sequence: 1 FLOLEPDAV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Published Applications\_AA.\*  
2: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US06\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US08\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	100.0	9	11	US-09-870-216C-5
2	44	100.0	9	13	US-10-017-327-5
3	37	84.1	1278	17	US-10-828-924-79
4	35	79.5	219	17	US-10-425-115-212999
5	35	79.5	461	17	US-10-425-115-213000
6	35	79.5	654	17	US-10-369-493-9433
7	35	79.5	680	14	US-10-437-963-17753
8	34	77.3	156	16	US-10-424-599-198066
9	33	75.0	125	15	US-10-369-493-12756
10	33	75.0	388	14	US-10-369-493-12756
11	32	72.7	225	15	US-10-282-122A-49767
12	32	72.7	274	15	US-10-282-122A-51999
13	32	72.7	414	14	US-10-260-937-5

14	32	72.7	414	14	US-10-260-937-49	Sequence 49, Appl
15	32	72.7	414	14	US-10-260-937-59	Sequence 59, Appl
16	32	72.7	551	15	US-10-382-248-8	Sequence 8, Appl1
17	32	72.7	551	17	US-10-741-853A-2	Sequence 2, Appl1
18	32	72.7	562	16	US-10-741-601-515	Sequence 515, App
19	32	72.7	580	14	US-10-260-937-34	Sequence 34, Appl
20	32	72.7	599	10	US-09-953-067A-3	Sequence 3, Appl1
21	32	72.7	599	14	US-10-097-340-264	Sequence 264, App
22	32	72.7	599	14	US-10-260-937-47	Sequence 47, Appl
23	32	72.7	599	15	US-10-382-248-6	Sequence 6, Appl1
24	32	72.7	599	16	US-10-741-601-514	Sequence 514, App
25	32	72.7	602	14	US-10-260-937-48	Sequence 48, Appl
26	32	72.7	602	10	US-09-953-067A-7	Sequence 7, Appl1
27	32	72.7	602	14	US-10-260-937-51	Sequence 51, Appl
28	32	72.7	602	14	US-10-260-937-52	Sequence 52, Appl
29	32	72.7	603	14	US-10-260-937-58	Sequence 58, Appl
30	32	72.7	606	14	US-10-260-937-50	Sequence 50, Appl
31	32	72.7	629	14	US-10-260-937-15	Sequence 15, Appl
32	32	72.7	630	17	US-10-783-237A-9	Sequence 9, Appl1
33	32	72.7	630	17	US-10-783-237A-11	Sequence 11, Appl1
34	32	72.7	632	14	US-10-260-937-60	Sequence 60, Appl
35	32	72.7	633	9	US-09-919-060-13	Sequence 13, Appl
36	32	72.7	633	15	US-10-260-937-2	Sequence 2, Appl1
37	32	72.7	633	14	US-10-678-140-13	Sequence 13, Appl1
38	32	72.7	641	14	US-10-369-493-8814	Sequence 8814, Ap
39	32	72.7	756	14	US-10-369-493-17121	Sequence 17121, A
40	31	70.5	141	16	US-10-437-963-160478	Sequence 160478,
41	31	70.5	167	16	US-10-437-963-194355	Sequence 194355,
42	31	70.5	188	14	US-10-308-460-4	Sequence 4, Appl1
43	31	70.5	190	14	US-10-308-460-2	Sequence 2, Appl1
44	31	70.5	190	14	US-10-308-460-6	Sequence 6, Appl1
45	31	70.5	195	15	US-10-276-774-2688	Sequence 2688, Ap

# ALIGNMENTS

RESULT 1  
US-09-870-216C-5  
; Sequence 5, Application US/09870216C  
; Publication No. US20040138135A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
; FILE REFERENCE: 68126881210100  
; CURRENT APPLICATION NUMBER: US/09/870, 216C  
; CURRENT FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/209, 391  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/226, 256  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: 60/257, 008  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-870-216C-5

Query Match 100.0%; Score 44; DB 11; Length 9;  
Best local Similarity 100.0%; Pred. No. 1.5e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
Db 1 FLOLEPDAV 9

RESULT 2  
US-10-017-327-5  
; Sequence 5, Application US/10017327

```

; Publication No. US20020155477A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Nicolette
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND
; TITLE OF INVENTION: METHODS FOR USING SAME
; FILE REFERENCE: G2 2101.20
; CURRENT APPLICATION NUMBER: US/10/017,327
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-017-327-5

Query Match
Best Local Similarity 100.0%; Score 44; DB 13; Length 9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9
Db 1 FLOLEPDAV 9

RESULT 3
US-10-828-924-79
; Sequence 79, Application US/10828924
; Publication No. US2005003401A1
; GENERAL INFORMATION:
; APPLICANT: STRATAGENE
; TITLE OF INVENTION: Pfu REPLICATION ACCESSORY FACTORS AND METHODS OF USE
; FILE REFERENCE: 04121.0161-00000
; CURRENT APPLICATION NUMBER: US/10/828,924
; CURRENT FILING DATE: 2004-04-20
; PRIOR APPLICATION NUMBER: 60/146,580
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Helicase 7
; US-10-828-924-79

Query Match
Best Local Similarity 84.1%; Score 37; DB 17; Length 1278;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPDA 8
Db 650 YLOLEPDA 657

RESULT 4
US-10-425-115-212999
; Sequence 212999, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 212999
; LENGTH: 219

```

```

; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_125856C.1.pep
; US-10-425-115-212999

Query Match
Best Local Similarity 79.5%; Score 35; DB 17; Length 219;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9
Db 72 FLOLEIDSV 80

RESULT 5
US-10-425-115-213000
; Sequence 213000, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 213000
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(461)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_125857C.1.pep
; US-10-425-115-213000

Query Match
Best Local Similarity 79.5%; Score 35; DB 17; Length 461;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9
Db 72 FLOLEIDSV 80

RESULT 6
US-10-369-493-9433
; Sequence 9433, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9433
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Xylella fastidiosa

```

US-10-369-493-9433

Query Match 79.5%; Score 35; DB 14; Length 654;  
Best Local Similarity 77.8%; Pred. No. 1.6e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
DB 89 FLOLEHDAV 97

RESULT 7

US-10-369-493-17753  
; Sequence 17753, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; PRIOR FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 17753  
; LENGTH: 680  
; TYPE: PRT  
; ORGANISM: Xylella fastidiosa  
US-10-369-493-17753

Query Match 79.5%; Score 35; DB 14; Length 680;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
DB 114 FLOLEHDAV 122

RESULT 8

US-10-437-963-198066  
; Sequence 198066, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; PRIOR FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 198066  
; LENGTH: 156  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_93763C.1.pep  
US-10-437-963-198066

Query Match 77.3%; Score 34; DB 16; Length 156;  
Best Local Similarity 100.0%; Pred. No. 58;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QLEPDAV 9  
DB 78 QLEPDAV 84

RESULT 9

US-10-424-599-274136  
; Sequence 274136, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; PRIOR FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 274136  
; LENGTH: 125  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(125)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_89567C.1.pep  
US-10-424-599-274136

Query Match 75.0%; Score 33; DB 15; Length 125;  
Best Local Similarity 77.8%; Pred. No. 74;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLEPDAV 9  
DB 77 FLOLEPDAV 85

RESULT 10

US-10-369-493-12756  
; Sequence 12756, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; PRIOR FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 12756  
; LENGTH: 388  
; TYPE: PRT  
; ORGANISM: Aspergillus nidulans  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(388)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-12756

Query Match 75.0%; Score 33; DB 14; Length 388;  
Best Local Similarity 87.5%; Pred. No. 2.4e+02;

```
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 LQLEFDAY 9
   ||| |||
Db 54 LQLEFDAY 61

RESULT 11
US-10-282-122A-49767
; Sequence 49767, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49767
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
US-10-282-122A-49767

Query Match 72.7%; Score 32; DB 15; Length 225;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 FLOLEFDAY 9
   ||| |||
Db 111 YLKLEFDV 119

RESULT 12
US-10-282-122A-51999
; Sequence 51999, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
```

```
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51999
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-10-282-122A-51999

Query Match 72.7%; Score 32; DB 15; Length 274;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FLOLEFDAY 9
   ||| |||
Db 205 FLKDFDGI 213

RESULT 13
US-10-260-937-5
; Sequence 5, Application US/10260937
; Publication No. US20030220306A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Daniel
; APPLICANT: Chandrasekharan, N. Vishvanath
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 07913-007001
; CURRENT APPLICATION NUMBER: US/10/260,937
; CURRENT FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: US 60/326,133
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/373,225
; PRIOR FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/373,661
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/411,575
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 89
```

SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 414  
; TYPE: PRF  
; ORGANISM: Canis familiaris  
US-10-260-937-5

Query Match  
Best Local Similarity 72.7%; Score 32; DB 14; Length 414;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLKD 7  
Db 170 FLOLKD 176

RESULT 14  
US-10-260-937-49  
; Sequence 49, Application US/10260937  
; Publication No. US20030220306A1  
; GENERAL INFORMATION:  
; APPLICANT: Simmons, Daniel  
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND  
; FILE REFERENCE: 07913-007001  
; CURRENT FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: US 60/326,133  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 60/373,225  
; PRIOR FILING DATE: 2002-04-15  
; PRIOR APPLICATION NUMBER: US 60/373,661  
; PRIOR FILING DATE: 2002-04-16  
; PRIOR APPLICATION NUMBER: US 60/411,575  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 49  
; LENGTH: 414  
; TYPE: PRF  
; ORGANISM: Canis familiaris  
US-10-260-937-49

Query Match  
Best Local Similarity 72.7%; Score 32; DB 14; Length 414;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLKD 7  
Db 170 FLOLKD 176

RESULT 15  
US-10-260-937-59  
; Sequence 59, Application US/10260937  
; Publication No. US20030220306A1  
; GENERAL INFORMATION:  
; APPLICANT: Simmons, Daniel  
; APPLICANT: Chandrasekharan, N. Vishvanath  
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND  
; FILE REFERENCE: 07913-007001  
; CURRENT FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: US 60/326,133  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 60/373,225  
; PRIOR FILING DATE: 2002-04-15  
; PRIOR APPLICATION NUMBER: US 60/373,661  
; PRIOR FILING DATE: 2002-04-16  
; PRIOR APPLICATION NUMBER: US 60/411,575  
; PRIOR FILING DATE: 2002-09-16

NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 59  
; LENGTH: 414  
; TYPE: PRF  
; ORGANISM: Canis familiaris  
US-10-260-937-59

Query Match  
Best Local Similarity 72.7%; Score 32; DB 14; Length 414;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLKD 7  
Db 170 FLOLKD 176

RESULT 16  
US-10-382-248-8  
; Sequence 8, Application US/10382248  
; Publication No. US20040058347A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook, et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-568C  
; CURRENT FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: US/10/382,248  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 60/366,928  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: 60/361,974  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: 60/365,477  
; PRIOR FILING DATE: 2002-03-19  
; PRIOR APPLICATION NUMBER: 60/401,661  
; PRIOR FILING DATE: 2002-08-06  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: Curaseqlast version 0.1  
; SEQ ID NO 8  
; LENGTH: 551  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-10-382-248-8

Query Match  
Best Local Similarity 72.7%; Score 32; DB 15; Length 551;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLKD 7  
Db 307 FLOLKD 313

RESULT 17  
US-10-741-853A-2  
; Sequence 2, Application US/10741853A  
; Publication No. US20040203030A1  
; GENERAL INFORMATION:  
; APPLICANT: Garrett-Engle, Philip, W  
; TITLE OF INVENTION: Alternatively Spliced Isoform of Human COX1  
; FILE REFERENCE: RS0206  
; CURRENT FILING DATE: 2003-12-18  
; PRIOR APPLICATION NUMBER: US 60/435,478  
; PRIOR FILING DATE: 2002-12-19  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 551  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-10-741-853A-2

Query Match  
Best Local Similarity 72.7%; Score 32; DB 17; Length 551;

Best Local Similarity 85.7%; Pred. No. 5.5e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEPFD 7  
|||:|  
Db 307 FLOLEKFD 313

RESULT 18  
US-10-741-601-515

; Sequence 515, Application US/10741601  
; Publication No. US20040166519A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; FILE REFERENCE: C0001500

; CURRENT APPLICATION NUMBER: US/10/741,601

; NUMBER OF SEQ ID NOS: 26415

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 515

; LENGTH: 562

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-741-601-515

Query Match 72.7%; Score 32; DB 16; Length 562;

Best Local Similarity 85.7%; Pred. No. 5.7e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEPFD 7  
|||:|  
Db 355 FLOLEKFD 361

RESULT 19  
US-10-260-937-34

; Sequence 34, Application US/10260937

; Publication No. US20030220306A1

; GENERAL INFORMATION:

; APPLICANT: Simmons, Daniel

; APPLICANT: Chandrasekharan, N. Vishvanath

; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND

; FILE REFERENCE: 07913-007001

; CURRENT APPLICATION NUMBER: US/10/260,937

; CURRENT FILING DATE: 2002-09-28

; PRIOR APPLICATION NUMBER: US 60/326,133

; PRIOR FILING DATE: 2001-09-28

; PRIOR APPLICATION NUMBER: US 60/373,225

; PRIOR FILING DATE: 2002-04-15

; PRIOR APPLICATION NUMBER: US 60/373,661

; PRIOR FILING DATE: 2002-04-16

; PRIOR APPLICATION NUMBER: US 60/411,575

; PRIOR FILING DATE: 2002-09-16

; NUMBER OF SEQ ID NOS: 89

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 34

; LENGTH: 580

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: consensus sequence

US-10-260-937-34

Query Match 72.7%; Score 32; DB 14; Length 580;

Best Local Similarity 85.7%; Pred. No. 5.8e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEPFD 7  
|||:|  
Db 344 FLOLEKFD 350

RESULT 20  
US-09-953-067A-3

; Sequence 3, Application US/09953067A

; Publication No. US20030082141A1

; GENERAL INFORMATION:

; APPLICANT: O'CONNOR, J. Patrick

; TITLE OF INVENTION: COX-2 FUNCTION AND WOUND HEALING

; FILE REFERENCE: 267/043

; CURRENT APPLICATION NUMBER: US/09/953,067A

; CURRENT FILING DATE: 2001-09-11

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 599

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-953-067A-3

Query Match 72.7%; Score 32; DB 10; Length 599;

Best Local Similarity 85.7%; Pred. No. 6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEPFD 7  
|||:|  
Db 355 FLOLEKFD 361

RESULT 21  
US-10-097-340-264

; Sequence 264, Application US/10097340

; Publication No. US20030087250A1

; GENERAL INFORMATION:

; APPLICANT: John MONAHAN

; APPLICANT: Manjula GANNAVAPU

; APPLICANT: Sebastian HOERSCH

; APPLICANT: Shubhangi KAMATYAR

; APPLICANT: Steve G. KOVATS

; APPLICANT: Michael E. MEYERS

; APPLICANT: Michael MORRISSEY

; APPLICANT: Peter OLANDT

; APPLICANT: Ami SEN

; APPLICANT: Peter VEIBY

; APPLICANT: Gordon B. MILLS

; APPLICANT: Robert C. BAST, Jr.

; APPLICANT: Karen LU

; APPLICANT: Rosemarie SCHMANDT

; APPLICANT: Xumei ZHAO

; APPLICANT: Karen GLATT

; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,

; FILE REFERENCE: MRI-030

; CURRENT APPLICATION NUMBER: US/10/097,340

; CURRENT FILING DATE: 2002-03-14

; PRIOR APPLICATION NUMBER: 60/276,025

; PRIOR FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 60/325,149

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: 60/276,026

; PRIOR FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 60/324,967

; PRIOR FILING DATE: 2001/09/26

; PRIOR APPLICATION NUMBER: 60/311,732

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: 60/325,102

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: 60/323,580

; PRIOR FILING DATE: 2001-09-19

; NUMBER OF SEQ ID NOS: 363

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 264

; LENGTH: 599

;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-097-340-264

Query Match  
Best Local Similarity 72.7%; Score 32; DB 14; Length 599;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFD 7  
|||:|  
Db 355 FLOLEFD 361

RESULT 22  
US-10-260-937-47  
; Sequence 47, Application US/10260937  
; Publication No. US20030220306A1  
; GENERAL INFORMATION:  
; APPLICANT: Simmons, Daniel  
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND  
; FILE REFERENCE: 07913-007001  
; CURRENT APPLICATION NUMBER: US/10/260,937  
; PRIOR FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: US 60/326,133  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 60/373,225  
; PRIOR FILING DATE: 2002-04-15  
; PRIOR APPLICATION NUMBER: US 60/373,661  
; PRIOR FILING DATE: 2002-04-16  
; PRIOR APPLICATION NUMBER: US 60/411,575  
; PRIOR FILING DATE: 2002-09-16  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 47  
; LENGTH: 599  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-260-937-47

Query Match  
Best Local Similarity 72.7%; Score 32; DB 14; Length 599;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFD 7  
|||:|  
Db 355 FLOLEFD 361

RESULT 23  
US-10-382-248-6  
; Sequence 6, Application US/10382248  
; Publication No. US20040058347A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook, et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-568C  
; CURRENT APPLICATION NUMBER: US/10/382,248  
; PRIOR FILING DATE: 2003-03-05  
; PRIOR APPLICATION NUMBER: 60/366,928  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 60/361,974  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: 60/365,477  
; PRIOR FILING DATE: 2002-03-19  
; PRIOR APPLICATION NUMBER: 60/401,661  
; PRIOR FILING DATE: 2002-08-06  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: Curaseqglst version 0.1  
; SEQ ID NO 6  
; LENGTH: 599  
; TYPE: PRT

;; ORGANISM: Homo sapiens  
US-10-382-248-6

Query Match  
Best Local Similarity 72.7%; Score 32; DB 15; Length 599;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFD 7  
|||:|  
Db 355 FLOLEFD 361

RESULT 24  
US-10-741-601-514  
; Sequence 514, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CLO01500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; PRIOR FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 514  
; LENGTH: 599  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-741-601-514

Query Match  
Best Local Similarity 72.7%; Score 32; DB 16; Length 599;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFD 7  
|||:|  
Db 355 FLOLEFD 361

RESULT 25  
US-10-260-937-48  
; Sequence 48, Application US/10260937  
; Publication No. US20030220306A1  
; GENERAL INFORMATION:  
; APPLICANT: Simmons, Daniel  
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND  
; FILE REFERENCE: 07913-007001  
; CURRENT APPLICATION NUMBER: US/10/260,937  
; PRIOR FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: US 60/326,133  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 60/373,225  
; PRIOR FILING DATE: 2002-04-15  
; PRIOR APPLICATION NUMBER: US 60/373,661  
; PRIOR FILING DATE: 2002-04-16  
; PRIOR APPLICATION NUMBER: US 60/411,575  
; PRIOR FILING DATE: 2002-09-16  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 48  
; LENGTH: 600  
; TYPE: PRT  
; ORGANISM: Ovis aries  
US-10-260-937-48

Query Match  
Best Local Similarity 72.7%; Score 32; DB 14; Length 600;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFD 7

Db 356 FLOLQKFD 362

RESULT 26  
US-09-953-067A-7  
; Sequence 7, Application US/09955067A  
; Publication No. US20030082141A1  
; GENERAL INFORMATION:  
; APPLICANT: O'CONNOR, J. Patrick  
; TITLE OF INVENTION: COX-2 FUNCTION AND WOUND HEALING  
; FILE REFERENCE: 267/043  
; CURRENT APPLICATION NUMBER: US/09/953,067A  
; CURRENT FILING DATE: 2001-09-11  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 602  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-953-067A-7

Query Match 72.7%; Score 32; DB 10; Length 602;  
Best Local Similarity 85.7%; Pred. No. 6.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLQKFD 7  
Db 358 FLOLQKFD 364

RESULT 27  
US-10-260-937-51  
; Sequence 51, Application US/10260937  
; Publication No. US20030220306A1  
; GENERAL INFORMATION:  
; APPLICANT: Simmons, Daniel  
; APPLICANT: Chandrasekharan, N. Vishvanath  
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND  
; FILE REFERENCE: 07913-007001  
; CURRENT APPLICATION NUMBER: US/10/260,937  
; CURRENT FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: US 60/326,133  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 60/373,225  
; PRIOR FILING DATE: 2002-04-15  
; PRIOR APPLICATION NUMBER: US 60/373,661  
; PRIOR FILING DATE: 2002-04-16  
; PRIOR APPLICATION NUMBER: US 60/411,575  
; PRIOR FILING DATE: 2002-09-16  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 51  
; LENGTH: 602  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-260-937-51

Query Match 72.7%; Score 32; DB 14; Length 602;  
Best Local Similarity 85.7%; Pred. No. 6.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLQKFD 7  
Db 358 FLOLQKFD 364

RESULT 28  
US-10-260-937-52  
; Sequence 52, Application US/10260937  
; Publication No. US20030220306A1  
; GENERAL INFORMATION:

APPLICANT: Simmons, Daniel  
; APPLICANT: Chandrasekharan, N. Vishvanath  
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND  
; FILE REFERENCE: 07913-007001  
; CURRENT APPLICATION NUMBER: US/10/260,937  
; CURRENT FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: US 60/326,133  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 60/373,225  
; PRIOR FILING DATE: 2002-04-15  
; PRIOR APPLICATION NUMBER: US 60/373,661  
; PRIOR FILING DATE: 2002-04-16  
; PRIOR APPLICATION NUMBER: US 60/411,575  
; PRIOR FILING DATE: 2002-09-16  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 52  
; LENGTH: 602  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-260-937-52

Query Match 72.7%; Score 32; DB 14; Length 602;  
Best Local Similarity 85.7%; Pred. No. 6.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLQKFD 7  
Db 358 FLOLQKFD 364

RESULT 29  
US-10-260-937-58  
; Sequence 58, Application US/10260937  
; Publication No. US20030220306A1  
; GENERAL INFORMATION:  
; APPLICANT: Simmons, Daniel  
; APPLICANT: Chandrasekharan, N. Vishvanath  
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND  
; FILE REFERENCE: 07913-007001  
; CURRENT APPLICATION NUMBER: US/10/260,937  
; CURRENT FILING DATE: 2002-09-28  
; PRIOR APPLICATION NUMBER: US 60/326,133  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US 60/373,225  
; PRIOR FILING DATE: 2002-04-15  
; PRIOR APPLICATION NUMBER: US 60/373,661  
; PRIOR FILING DATE: 2002-04-16  
; PRIOR APPLICATION NUMBER: US 60/411,575  
; PRIOR FILING DATE: 2002-09-16  
; NUMBER OF SEQ ID NOS: 89  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 58  
; LENGTH: 603  
; TYPE: PRT  
; ORGANISM: Canis familiaris  
US-10-260-937-58

Query Match 72.7%; Score 32; DB 14; Length 603;  
Best Local Similarity 85.7%; Pred. No. 6.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLQKFD 7  
Db 359 FLOLQKFD 365

RESULT 30  
US-10-260-937-50  
; Sequence 50, Application US/10260937  
; Publication No. US20030220306A1



```

; GENERAL INFORMATION:
; APPLICANT: Simmons, Daniel
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND
; FILE REFERENCE: 07913-007001
; CURRENT FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: US 60/326,133
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/373,225
; PRIOR FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/373,661
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/411,575
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-10-260-937-50

```

```

Query Match      72.7%; Score 32; DB 14; Length 606;
Best Local Similarity 85.7%; Pred. No. 6.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FLOLEPD 7
      |||||:
Db      362 FLOLKFD 368

```

```

RESULT 31
US-10-260-937-15
; Sequence 15, Application US/10260937
; Publication No. US20030220306A1
; GENERAL INFORMATION:
; APPLICANT: Chandrasekharan, N. Vishvanath
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND
; FILE REFERENCE: 07913-007001
; CURRENT FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: US 60/326,133
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/373,225
; PRIOR FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/373,661
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/411,575
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 629
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-260-937-15

```

```

Query Match      72.7%; Score 32; DB 14; Length 629;
Best Local Similarity 85.7%; Pred. No. 6.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FLOLEPD 7
      |||||:
Db      385 FLOLKFD 391

```

```

RESULT 32
US-10-783-297A-9
; Sequence 9, Application US/10783297A

```

```

; Publication No. US20040235017A1
; GENERAL INFORMATION:
; APPLICANT: Johnson & Johnson Pharmaceutical Research and development
; TITLE OF INVENTION: HUMAN CYCLOOXYGENASE-3 AND USES THEREOF
; FILE REFERENCE: PRD-2041
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/783,297A
; PRIOR FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-783-297A-9

```

```

Query Match      72.7%; Score 32; DB 17; Length 630;
Best Local Similarity 85.7%; Pred. No. 6.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FLOLEPD 7
      |||||:
Db      386 FLOLKFD 392

```

```

RESULT 33
US-10-783-297A-11
; Sequence 11, Application US/10783297A
; Publication No. US20040235017A1
; GENERAL INFORMATION:
; APPLICANT: Johnson & Johnson Pharmaceutical Research and development
; TITLE OF INVENTION: HUMAN CYCLOOXYGENASE-3 AND USES THEREOF
; FILE REFERENCE: PRD-2041
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/10/783,297A
; PRIOR FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-783-297A-11

```

```

Query Match      72.7%; Score 32; DB 17; Length 630;
Best Local Similarity 85.7%; Pred. No. 6.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FLOLEPD 7
      |||||:
Db      386 FLOLKFD 392

```

```

RESULT 34
US-10-260-937-60
; Sequence 60, Application US/10260937
; Publication No. US20030220306A1
; GENERAL INFORMATION:
; APPLICANT: Chandrasekharan, N. Vishvanath
; TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND
; FILE REFERENCE: 07913-007001
; CURRENT FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: US/10/260,937
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: US 60/326,133
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/373,225
; PRIOR FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/373,661
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/411,575
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 4.0
US-10-260-937-60

```

SEQ ID NO 60  
LENGTH: 632  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: consensus sequence  
US-10-260-937-60

Query Match 72.7%; Score 32; DB 14; Length 632;  
Best Local Similarity 85.7%; Pred. No. 6.4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPD 7  
DB 389 FLOLKEPD 395

RESULT 35  
US-09-919-060-13  
Sequence 13, Application US/09919060  
Patent No. US20020064845A1  
GENERAL INFORMATION:  
APPLICANT: Wisniewski, Nancy  
APPLICANT: Brandt, Kevin S.  
TITLE OF INVENTION: CANINE COX-1 AND COX-2 NUCLEIC ACID MOLECULES, PROTEINS AND USES  
FILE REFERENCE: AD-1  
CURRENT APPLICATION NUMBER: US/09/919,060  
CURRENT FILING DATE: 2001-07-31  
PRIOR APPLICATION NUMBER: 60/224,486  
PRIOR FILING DATE: 2000-08-11  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 13  
LENGTH: 633  
TYPE: PRT  
ORGANISM: Canis familiaris  
US-09-919-060-13

Query Match 72.7%; Score 32; DB 9; Length 633;  
Best Local Similarity 85.7%; Pred. No. 6.4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPD 7  
DB 389 FLOLKEPD 395

RESULT 36  
US-10-260-937-2  
Sequence 2, Application US/10260937  
Publication No. US20030220306A1  
GENERAL INFORMATION:  
APPLICANT: Simmons, Daniel  
APPLICANT: Chandrasekharan, N. Vishvanath  
TITLE OF INVENTION: NOVEL CYCLOOXYGENASE VARIANTS AND  
FILE REFERENCE: 07913-007001  
CURRENT APPLICATION NUMBER: US/10/260,937  
CURRENT FILING DATE: 2002-09-28  
PRIOR APPLICATION NUMBER: US 60/326,133  
PRIOR FILING DATE: 2001-09-28  
PRIOR APPLICATION NUMBER: US 60/373,225  
PRIOR FILING DATE: 2002-04-15  
PRIOR APPLICATION NUMBER: US 60/373,661  
PRIOR FILING DATE: 2002-04-16  
PRIOR APPLICATION NUMBER: US 60/411,575  
PRIOR FILING DATE: 2002-09-16  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 633  
TYPE: PRT  
ORGANISM: Canis familiaris

US-10-260-937-2

Query Match 72.7%; Score 32; DB 14; Length 633;  
Best Local Similarity 85.7%; Pred. No. 6.4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPD 7  
DB 389 FLOLKEPD 395

RESULT 37  
US-10-679-140-13  
Sequence 13, Application US/10679140  
Publication No. US20040091986A1  
GENERAL INFORMATION:  
APPLICANT: Wisniewski, Nancy  
APPLICANT: Brandt, Kevin S.  
TITLE OF INVENTION: CANINE COX-1 AND COX-2 NUCLEIC ACID MOLECULES, PROTEINS AND USES  
FILE REFERENCE: AD-1  
CURRENT APPLICATION NUMBER: US/10/679,140  
CURRENT FILING DATE: 2003-10-22  
PRIOR APPLICATION NUMBER: US/09/919,060  
PRIOR FILING DATE: 2001-07-31  
PRIOR APPLICATION NUMBER: 60/224,486  
PRIOR FILING DATE: 2000-08-11  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 13  
LENGTH: 633  
TYPE: PRT  
ORGANISM: Canis familiaris  
US-10-679-140-13

Query Match 72.7%; Score 32; DB 15; Length 633;  
Best Local Similarity 85.7%; Pred. No. 6.4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPD 7  
DB 389 FLOLKEPD 395

RESULT 38  
US-10-369-493-8814  
Sequence 8814, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 8814  
LENGTH: 641  
TYPE: PRT  
ORGANISM: Ralstonia metallidurans  
US-10-369-493-8814

Query Match 72.7%; Score 32; DB 14; Length 641;  
Best Local Similarity 55.6%; Pred. No. 6.5e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDV 9  
DB 389 FLOLKEPDV 395

Db 512 FINIDFNAV 520

RESULT 39  
US-10-369-493-17121

; Sequence 17121, Application US/10369493  
; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 17121

; LENGTH: 756

; TYPE: PRT

; ORGANISM: Bacillus halodurans

; US-10-369-493-17121

Query Match 72.7%; Score 32; DB 14; Length 756;  
Best Local Similarity 77.8%; Pred. No. 7.7e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLEFNAV 9

Db 235 FLOTFYFNAV 243

RESULT 40  
US-10-437-963-160478

; Sequence 160478, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovacic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Mu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 160478

; LENGTH: 141

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(141)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_59753C.1.pep

; US-10-437-963-160478

Query Match 70.5%; Score 31; DB 16; Length 141;  
Best Local Similarity 85.7%; Pred. No. 2.1e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEFNAV 7

Db 25 FLOLRFD 31

Search completed: January 12, 2005, 20:48:25  
Job time : 95 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:46:14 ; Search time 16.4 Seconds  
(without alignments)  
52.802 Million cell updates/sec

Title: US-09-870-216C-5  
Perfect score: 44  
Sequence: 1 FLOLEPDV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	79.5	680	2 G82526	ATP-dependent DNA
2	33	75.0	430	2 B82096	hypothetical prote
3	33	75.0	430	2 B82096	conserved hypotnet
4	32	72.7	154	2 G95328	hypothetical prote
5	32	72.7	228	2 T12000	cytochrome-c oxida
6	32	72.7	228	2 T09802	cytochrome-c oxida
7	32	72.7	229	2 S50328	cytochrome-c oxida
8	32	72.7	274	2 G97034	probable metal-dep
9	32	72.7	599	2 JH0259	prostaglandin-endo
10	32	72.7	599	2 A29947	prostaglandin-endo
11	32	72.7	600	2 S00561	prostaglandin-endo
12	32	72.7	600	2 A28960	prostaglandin-endo
13	32	72.7	602	2 S39782	cytochrome-c oxida
14	32	72.7	602	2 A35564	prostaglandin-endo
15	32	72.7	602	2 S69198	prostaglandin G/H
16	32	72.7	756	2 P83704	homocysteine methyl
17	31	70.5	204	2 A32252	probable phosphoe
18	31	70.5	210	2 A40979	temperature shock-
19	31	70.5	321	2 G70415	nucleotide sugar e
20	31	70.5	360	2 T06786	6a-hydroxymackia
21	31	70.5	384	2 A86324	protein P14D16.19
22	31	70.5	390	2 P81393	probable altronate
23	31	70.5	400	2 B90139	phosphomethylpyrim
24	31	70.5	578	2 T44444	hypothetical prote
25	31	70.5	578	2 T19678	hypothetical prote
26	31	70.5	1042	2 G64514	type I restriction
27	31	70.5	1852	2 JCS546	chitin synthase (B
28	30	68.2	100	2 A52073	hypothetical prote
29	30	68.2	155	2 B84198	hypothetical prote

30	30	68.2	211	2 D69888	micrococcal nuclea
31	30	68.2	226	2 G45170	cytochrome-c oxida
32	30	68.2	229	2 G90730	antitermination pr
33	30	68.2	237	2 B81870	probable membrane
34	30	68.2	257	2 D85581	hypothetical prote
35	30	68.2	272	1 G82076	3',5'-cyclic-nucle
36	30	68.2	280	2 T38816	hypothetical prote
37	30	68.2	290	2 B97100	pyridoxal kinase r
38	30	68.2	296	2 T12770	probable endonucle
39	30	68.2	299	2 T23932	hypothetical prote
40	30	68.2	300	2 C69857	formyltetrahydrofo
41	30	68.2	301	2 C85574	probable LysK-like
42	30	68.2	301	2 C90723	probable transcript
43	30	68.2	365	2 T04247	hypothetical prote
44	30	68.2	372	2 T01551	receptor kinase ho
45	30	68.2	420	2 T51874	hypothetical prote

ALIGNMENTS

RESULT 1  
G82526  
ATP-dependent DNA helicase XF2680 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: G82526  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: G82526  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-680 <STM>  
A:Cross-references: UNIPROT:Q9PA40; GB:A8004074; GB:A8003849; NID:G9107918; PIDN:AAFB547  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Brienes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrer, D.M.; Carrer, H  
as-Neto, E.; Docena, C.; El-Dorri, H.; Falcinani, A.F.; Ferreira, A.J.S.  
Submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kltajima, J.P.; Klieger, J.B.; Kuxmae, E.E.; Laigr  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marinho, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.U. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2680  
C:Superfamily: helicase II  
Query Match 79.5%; Score 35; DB 2; Length 680;  
Best Local Similarity 77.8%; Pred. No. 22;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLOLEPDV 9  
DB 114 FLOLEPDV 122  
RESULT 2  
G64232  
hypothetical protein MG296 - Mycoplasma genitalium  
C:Species: Mycoplasma genitalium  
C:Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: G64232  
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;  
M.; Fuhrmann, J.; Nguyen, D.; Ucterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.N.

C.A.; Venter, J.C.  
 Science 270, 397-403, 1995  
 A>Title: The minimal gene complement of *Mycoplasma genitalium*.  
 A/Reference number: A64200; MUID:96026346; PMID:7569993  
 A/Accession: G64232  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-129 <TIGR>  
 A/Cross-references: UNIPROT:P47538; GB:U39710; GB:U43967; NID:gl045989; PID:gl045994; TI  
 A/Experimental source: strain G-37  
 C/Genetics:  
 A/Genetic code: SGC3

Query Match 75.0%; Score 33; DB 2; Length 129;  
 Best Local Similarity 77.8%; Pred. No. 9;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLEFPAV 9  
 |||||  
 DB 11 FLOLEFPAV 19

RESULT 3  
 882096  
 conserved hypothetical protein VC2278 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)  
 C/Species: *Vibrio cholerae*  
 C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C/Accession: B82096  
 R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.;  
 charlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.  
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 A/Reference number: A82035; MUID:20406833; PMID:10952201  
 A/Accession: B82096  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-430 <HEI>  
 A/Cross-references: UNIPROT:Q9KPT4; GB:AE004299; GB:AE003852; NID:g9656835; PIDN:AAF9542  
 A/Experimental source: serogroup O1, strain N16961; biotype El Tor  
 C/Genetics:  
 A/Map position: 1  
 C/Superfamily: conserved hypothetical protein HI0125

Query Match 75.0%; Score 33; DB 2; Length 430;  
 Best Local Similarity 66.7%; Pred. No. 35;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEFPAV 9  
 |||||  
 DB 227 FMOUDFSAY 235

RESULT 4  
 G95328  
 hypothetical protein SMO0994 [imported] - *Sinorhizobium meliloti* (strain 1021) magaplaem  
 C/Species: *Sinorhizobium meliloti*  
 C/Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
 C/Accession: G95328  
 R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe  
 ; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Weller, D.H.; Yeh, K.C.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
 A>Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*  
 A/Reference number: A95262; MUID:21396509; PMID:11481432  
 A/Accession: G95328  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-154 <KUR>  
 A/Cross-references: UNIPROT:Q9Z2F1; GB:AE006469; PIDN:AAK5193.1; PID:gl4523639; GSPDB:C  
 A/Experimental source: strain 1021, megaplasmid pSymA  
 R/Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
 F.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A/Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 hebaunt, P.; Vandenbol, M.; Vorholter, F.J.; Welter, S.; Weller, D.H.; Wong, K.; Yeh, K.C.  
 A>Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
 A/Reference number: A96039; MUID:21368234; PMID:11474104  
 A/Contents: annotation  
 C/Genetics:  
 A/Genome: SMO0994  
 A/Genome: plasmid

Query Match 72.7%; Score 32; DB 2; Length 154;  
 Best Local Similarity 87.5%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FLOLEFPAV 9  
 |||||  
 DB 109 FLOLEFPAV 116

RESULT 5  
 T12000  
 cytochrome-c oxidase (EC 1.9.3.1) chain II - *Anopheles quadrimaculatus* A mitochondrion  
 C/Species: *Anopheles quadrimaculatus* A  
 C/Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 18-Aug-2003  
 C/Accession: T12000  
 R/Cockburn, A.F.; Mitchell, S.E.; Sawright, J.A.  
 Arch. Insect Biochem. Physiol. 14, 31-36, 1990  
 A>Title: Cloning of the mitochondrial genome of *Anopheles quadrimaculatus*.  
 A/Reference number: Z17375; MUID:92190510; PMID:2134168  
 A/Accession: T12000  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-228 <COC>  
 A/Cross-references: EMBL:U04272; NID:g342501; PID:g507281; PIDN:AAA93542.1  
 A/Experimental source: strain Orlando  
 C/Genetics:  
 A/Genome: mitochondrion  
 A/Note: COXII  
 C/Superfamily: cytochrome-c oxidase chain II, mitochondrial type; cytochrome-c oxidase c  
 ;Keywords: copper; electron transfer; membrane-associated complex; mitochondrial inner  
 F;161,196,200,207/Binding site: copper 1 (His, Cys, Cys, Met) #status predicted  
 F;161,196,200,204/Binding site: copper 2 (Cys, Glu, Cys, His) #status predicted  
 F;198/Binding site: magnesium (Glu) (shared with chain I) #status predicted

Query Match 72.7%; Score 32; DB 2; Length 228;  
 Best Local Similarity 75.0%; Pred. No. 28;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEFPAV 8  
 |||||  
 DB 113 FLOLEFPAV 120

RESULT 6  
 T09802  
 cytochrome-c oxidase (EC 1.9.3.1) chain II - *African malaria mosquito* mitochondrion  
 C/Species: *Anopheles gambiae* (African malaria mosquito)  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: T09802  
 R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe  
 ; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Weller, D.H.; Yeh, K.C.  
 Insect Mol. Biol. 2, 103-114, 1993  
 A>Title: The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genom  
 A/Reference number: Z16863; MUID:97242550; PMID:9087549  
 A/Accession: T09802  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-228 <BEA>  
 A/Cross-references: UNIPROT:P34840; EMBL:L20934; NID:g309056; PIDN:AD12192.1; PID:g3090  
 A/Experimental source: strain G3  
 C/Genetics:  
 A/Genome: mitochondrion  
 A/Genetic code: SGC4

C:Superfamily: cytochrome-c oxidase chain II, mitochondrial type; cytochrome-c oxidase c  
C:Keywords: copper; electron transfer; membrane-associated complex; mitochondrial inner  
F:9-214/Domain: cytochrome-c oxidase chain II homology <CO2>  
F:161,196,200,207/Binding site: copper 1 (His, Cys, Cys, Met) #status predicted  
F:196,198,200,204/Binding site: copper 2 (Cys, Glu, Cys, His) #status predicted  
F:198/Binding site: magnesium (Glu) (shared with chain I) #status predicted

Query Match 72.7% Score 32; DB 2; Length 228;  
Best Local Similarity 75.0%; Pred. No. 28;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Db 113 FLNLEPDS 120

RESULT 7  
S50328  
Cytochrome-c oxidase (EC 1.9.3.1) chain II - Katharina tunicata mitochondrion  
C:Species: mitochondrion Katharina tunicata  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S50328  
R:Boore, J.L.; Brown, W.M.  
Genetic 138, 423-443, 1394  
A:Title: Complete DNA sequence of the mitochondrial genome of the black chiton, Katharin  
A:Reference number: S50327; MUID:95129806; PMID:7828825  
A:Accession: S50328  
A:Molecule type: DNA  
A:Residues: 1-229 <BO>  
A:Cross-references: UNIPROT:O37534; EMBL:U09810; NID:9557273; PID:AA048365.1; PID:95572  
C:Genetics:  
A:Genome: mitochondrion  
A:Genetic code: SGC4  
C:Superfamily: cytochrome-c oxidase chain II, mitochondrial type; cytochrome-c oxidase c  
C:Keywords: copper; electron transfer; heme; membrane-associated complex; mitochondrial  
ein  
F:9-214/Domain: cytochrome-c oxidase chain II homology <CO2>  
F:161,196,200,207/Binding site: copper 1 (His, Cys, Cys, Met) #status predicted  
F:196,198,200,204/Binding site: copper 2 (Cys, Glu, Cys, His) #status predicted  
F:198/Binding site: magnesium (Glu) (shared with chain I) #status predicted

Query Match 72.7% Score 32; DB 2; Length 229;  
Best Local Similarity 75.0%; Pred. No. 28;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Cy 1 FLNLEPDA 8  
Db 113 FLNLEPDS 120

RESULT 8  
G97034  
Probable metal-dependent phosphoserine (PHP family), YcIV ortholog [imported] - Clostr  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: G97034  
R:Nolling, J.; Breston, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Kohn, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: G97034  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-274 <KUR>  
A:Cross-references: UNIPROT:Q97K29; GB:AE001437; PIDN:AAK79066.1; PID:915024008; GSPDB:C  
A:Genetics:  
A:Experimental source: Clostridium acetobutylicum ATCC824  
A:Gene: CAC1092  
C:Superfamily: hypothetical protein H11400

Query Match 72.7% Score 32; DB 2; Length 274;  
Best Local Similarity 55.6%; Pred. No. 34;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
Cy 1 FLNLEPDAV 9  
Db 205 FLNLEPDS 213

RESULT 9  
JH0259  
Prostaglandin-endoperoxide synthase (EC 1.14.99.1) 1 precursor - human  
N:Alternate names: cyclooxygenase; prostaglandin G/H synthase; prostaglandin H synthase  
N:Contains: prostaglandin-endoperoxide synthase 1, splice form 2  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C:Accession: JH0259; PH0225; A3937; B38146; A38146; S50181; A36746; S69169  
R:Takahashi, Y.; Ueda, N.; Yoshimoto, T.; Yamamoto, S.; Yokoyama, C.; Miyata, A.; Tanabe  
Biochem. Biophys. Res. Commun. 182, 433-438, 1992  
A:Title: Immunodominant purification and cDNA cloning of human platelet prostaglandin en  
A:Reference number: JH0259; MUID:92134251; PMID:1734857  
A:Accession: JH0259  
A:Molecule type: mRNA  
A:Residues: 1-599 <TKL>  
A:Cross-references: UNIPROT:P23219; GB:S78220; NID:9243971; PIDN:AA021215.1; PID:9243972  
A:Accession: PH0225  
A:Molecule type: protein  
A:Residues: 24-31 <TKX2>  
R:Funk, C.D.; Funk, L.B.; Kennedy, M.E.; Pong, A.S.; Fitzgerald, G.A.  
FASEB J. 5, 2304-2312, 1991  
A:Title: Human platelet/erythrocyte leukemia cell prostaglandin G/H synthase: cDNA cloning, e  
A:Reference number: A3937; MUID:91317397; PMID:1907252  
A:Accession: A3937  
A:Molecule type: mRNA  
A:Residues: 1-599 <PUN>  
A:Cross-references: GB:M5979; NID:9189886; PIDN:AAA03630.1; PID:9189887  
R:DiSiz, A.; Reginato, A.M.; Jimenez, S.A.  
J. Biol. Chem. 267, 10816-10822, 1992  
A:Title: Alternative splicing of human prostaglandin G/H synthase mRNA and evidence of d  
a tumor necrosis factor alpha.  
A:Reference number: A38146; MUID:92268138; PMID:1587858  
A:Accession: B38146  
A:Molecule type: mRNA  
A:Residues: 1-599 <DIA1>  
A:Cross-references: GB:S36271; NID:9249625; PIDN:AA022217.1; PID:9249626  
A:Experimental source: lung fibroblast  
A>Note: sequence extracted from NCBI backbone (NCBIN:103945, NCBI:P:103946)  
A:Accession: A38146  
A:Molecule type: mRNA  
A:Residues: 1-395, 433-599 <DIA2>  
A:Cross-references: GB:S36219; NID:9249623; PIDN:AA022216.1; PID:9249624  
A:Experimental source: lung fibroblast  
A>Note: sequence extracted from NCBI backbone (NCBIN:103825, NCBI:P:103826)  
R:Barnett, U.; Chow, J.; Ives, D.; Chiu, M.; Mackenzie, R.; Owen, E.; Nguyen, B.; Tsing  
Biochim. Biophys. Acta 1209, 130-139, 1994  
A:Title: Purification, characterization and selective inhibition of human prostaglandin  
A:Reference number: S50181; MUID:95035046; PMID:7947975  
A:Accession: S50181  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 24-32 <BAX>  
R:Yokoyama, C.; Tanabe, T.  
Biochem. Biophys. Res. Commun. 165, 888-894, 1989  
A:Title: Cloning of human gene encoding prostaglandin endoperoxide synthase and primary  
A:Reference number: A36746; MUID:90088508; PMID:2512924  
A:Accession: A36746  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-11, 'L', '13-112, 'L', '114-377, 'T', '379-599 <YOK>  
A:Cross-references: GB:M1822; NID:9189898; PIDN:AAA3639.1; PID:9387018  
R:Ren, Y.; Loose-Mitchell, D.S.; Kulmacz, R.J.  
Arch. Biochem. Biophys. 316, 751-757, 1995  
A:Title: Prostaglandin H synthase-1: evaluation of C-terminus function.  
A:Reference number: S69169; MUID:95168861; PMID:7864630  
A:Accession: S69169

A:Molecule type: protein  
 A:Residues: 585-599 <REN>  
 C:Genetics:  
 A:Gene: GDB:PTGS1  
 A:Cross-references: GDB:128070; OMIM:176805  
 A:Map position: 9q32-9q33.3  
 C:Function:  
 A:Description: catalyzes the oxidative cyclization by oxygen of arachidonic acid to prostaglandin biosynthesis  
 A:Pathway: prostaglandin biosynthesis  
 C:Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
 C:Keywords: alternative splicing; chromoprotein; endoplasmic reticulum; glycoprotein; he  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-599/Product: prostaglandin-endoperoxide synthase 1 #status experimental <MAT>  
 F:35-68/Domain: EGF homology <EGF>  
 F:103,143,409/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:308/Binding site: heme iron (His) (axial ligand) #status predicted  
 F:384,529/Active site: Tyr, Ser #status predicted

Query Match 72.7%; Score 32; DB 2; Length 599;  
 Best Local Similarity 85.7%; Pred. No. 83;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLQEPD 7  
 |||||  
 Db 355 FLOLQEPD 361

RESULT 10  
 A29947  
 prostaglandin-endoperoxide synthase (EC 1.14.99.1) precursor - sheep  
 N:Alternate names: prostaglandin endoperoxide synthetase  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 09-Jul-2004  
 C:Accession: A29947  
 R:Merlie, J.P.; Fagan, D.; Mudd, J.; Needleman, P.  
 J. Biol. Chem. 263, 3550-3553, 1988  
 A:Title: Isolation and characterization of the complementary DNA for sheep seminal vesic  
 A:Reference number: A29947; PMID:8153641; PMID:2831188  
 A:Accession: A29947  
 A:Molecule type: mRNA  
 A:Residues: 1-599 <MER>  
 A:Cross-references: UNIPROT:P05979; GB:M18243; NID:G165843; PIDN:AAA31511.1; PID:G165844  
 C:Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
 C:Keywords: oxidoreductase  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-599/Product: prostaglandin-endoperoxide synthase #status predicted <MAT>  
 F:35-68/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 599;  
 Best Local Similarity 85.7%; Pred. No. 83;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLQEPD 7  
 |||||  
 Db 355 FLOLQEPD 361

RESULT 11  
 S00561  
 prostaglandin-endoperoxide synthase (EC 1.14.99.1) precursor - sheep  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 10-Dec-1999  
 C:Accession: S00561  
 R:Yokoyama, C.; Takai, T.; Tanabe, T.  
 FEBS Lett. 231, 347-351, 1988  
 A:Title: Primary structure of sheep prostaglandin endoperoxide synthase deduced from cDN  
 A:Reference number: S00561; PMID:88196421; PMID:3129310  
 A:Accession: S00561  
 A:Molecule type: mRNA  
 A:Residues: 1-600 <YOK>  
 A:Cross-references: EMBL:Y00750; NID:G1361; PIDN:CAA68719.1; PID:G1362  
 A:Note: Part of this sequence, including the amino end of the mature protein, was confir

A:Note: 97-His, 164-Gly, 456-Gln, 520-Gln, 520-Lys, and 525-Ile were also found  
 C:Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
 C:Keywords: oxidoreductase  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-600/Product: prostaglandin-endoperoxide synthase #status experimental <MAT>  
 F:36-69/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 600;  
 Best Local Similarity 85.7%; Pred. No. 83;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLQEPD 7  
 |||||  
 Db 356 FLOLQEPD 362

RESULT 12  
 A28960  
 prostaglandin-endoperoxide synthase (EC 1.14.99.1) G/H precursor - sheep  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 09-Jul-2004  
 C:Accession: A28960  
 R:Dewitt, D.L.; Smith, W.L.  
 Proc. Natl. Acad. Sci. U.S.A. 85, 1412-1416, 1988  
 A:Title: Primary structure of prostaglandin G/H synthase from sheep vesicular gland dete  
 A:Reference number: A28960; PMID:8144447; PMID:3125548  
 A:Accession: A28960  
 A:Molecule type: mRNA  
 A:Residues: 1-600 <DEM>  
 A:Cross-references: UNIPROT:P05979; GB:J03599; NID:G166035; PIDN:AAA31576.1; PID:G166036  
 C:Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
 C:Keywords: oxidoreductase  
 F:36-69/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 600;  
 Best Local Similarity 85.7%; Pred. No. 83;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLQEPD 7  
 |||||  
 Db 356 FLOLQEPD 362

RESULT 13  
 S39782  
 cyclooxygenase 1 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 20-May-1994 #sequence\_revision 10-Nov-1995 #text\_change 10-Dec-1999  
 C:Accession: S39782  
 R:Feng, L.; Sun, W.; Xia, Y.; Tang, W.W.; Channugam, P.; Soyoola, E.; Wilson, C.B.; Hwan  
 Arch. Biochem. Biophys. 307, 361-368, 1993  
 A:Title: Cloning two isoforms of rat cyclooxygenase: differential regulation of their ex  
 A:Reference number: S39782; PMID:9409619; PMID:8274023  
 A:Accession: S39782  
 A:Molecule type: mRNA  
 A:Residues: 1-602 <FEN>  
 A:Cross-references: GB:S67721; NID:9460555; PIDN:AA829400.1; PID:9460556  
 C:Superfamily: human prostaglandin-endoperoxide synthase; EGF homology  
 F:38-71/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 602;  
 Best Local Similarity 85.7%; Pred. No. 83;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLQEPD 7  
 |||||  
 Db 358 FLOLQEPD 364

RESULT 14  
 A3564  
 prostaglandin-endoperoxide synthase (EC 1.14.99.1) precursor - mouse  
 C:Species: Mus musculus (house mouse)



C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 09-Jul-2004  
 C/Accession: A35564  
 C/Dewitt, D.L.; El-Harith, E.A.; Kraemer, S.A.; Andrews, M.J.; Yao, E.F.; Armstrong, R.I.  
 J. Biol. Chem. 265, 5192-5198, 1990  
 A>Title: The aspartin and heme-binding sites of ovine and murine prostaglandin endoperox-  
 A/Reference number: A35564; MUID:90203007; PMID:2108169  
 A/Accession: A35564  
 A/Molecule type: mRNA  
 A/Residues: 1-602 <DEM>  
 A/Cross-references: UNIPROT:P22437; GB:M34141; NID:G200302; PIDN:AAA39913.1; PID:G200303  
 C/Species: human prostaglandin-endoperoxide synthase; EGF homology  
 C/Keywords: oxidoreductase  
 F:1-26/Domain: signal sequence #status predicted <SIG>  
 F:27-602/Product: prostaglandin-endoperoxide synthase #status predicted <MAY>  
 F:38-71/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 602;  
 Best Local Similarity 85.7%; Pred. No. 83;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEPFD 7  
 |||||  
 Db 358 FLOQLKFD 364

RESULT 15  
 S69198  
 prostaglandin G/H synthase 1 - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C>Date: 24-Aug-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C/Accession: S69198; S69199  
 R/Kitzler, J.W.  
 submitted to the EMBL Data Library, December 1994  
 A/Reference number: S69198  
 A/Accession: S69198  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-602 <KIT>  
 A/Cross-references: UNIPROT:Q63921; EMBL:U18060; NID:G603051; PIDN:AAA85823.1; PID:G6030  
 R/Kitzler, J.; Hill, E.; Hardman, R.; Reddy, N.; Philippot, R.; Elling, T.E.  
 Arch. Biochem. Biophys. 316, 856-863, 1995  
 A>Title: Analysis and quantitation of splicing variants of the TPA-inducible PGHS-1 mRNA  
 A/Reference number: S69199; MUID:95168876; PMID:7864644  
 A/Accession: S69199  
 A/Molecule type: mRNA  
 A/Residues: 61-602 <KIT>  
 A/Cross-references: EMBL:U18060  
 A/Note: only a part of the nucleic acid sequence is shown  
 C/Species: human prostaglandin-endoperoxide synthase; EGF homology  
 C/Keywords: alternative splicing  
 F:38-71/Domain: EGF homology <EGF>

Query Match 72.7%; Score 32; DB 2; Length 602;  
 Best Local Similarity 85.7%; Pred. No. 83;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEPFD 7  
 |||||  
 Db 358 FLOQLKFD 364

RESULT 16  
 F83704  
 homoyserin methyl transferase BH0438 [imported] - Bacillus halodurans (strain C-125)  
 C/Species: Bacillus halodurans  
 C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C/Accession: F83704  
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A/Reference number: A83650; MUID:20512582; PMID:11058132  
 A/Accession: F83704  
 A/Status: preliminary

A/Molecule type: DNA  
 A/Residues: 1-756 <STO>  
 A/Cross-references: UNIPROT:Q9KFP1; GB:AP001508; GB:BA000004; NID:G10172890; PIDN:BA8041  
 A/Experimental source: strain C-125  
 C/Genetics:  
 A/Gene: BH0438  
 C/Superfamily: cobalamin-independent methionine synthase

Query Match 72.7%; Score 32; DB 2; Length 756;  
 Best Local Similarity 77.8%; Pred. No. 1,1e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLOLEPDAV 9  
 |||||  
 Db 235 FLOQYDAV 243

RESULT 17  
 A32252  
 probable phosphoesterase (EC 3.1.1.-) L - Acinetobacter calcoaceticus (fragment)  
 N/Alternate names: gene L protein  
 C/Species: Acinetobacter calcoaceticus  
 C/Date: 12-Oct-1989 #sequence\_revision 12-Oct-1989 #text\_change 09-Jul-2004  
 C/Accession: A32252; T01631  
 R/Goosen, N.; Horeman, H.P.A.; Huinen, R.G.M.; van de Putte, P.  
 J. Bacteriol. 171, 447-455, 1989  
 A>Title: Acinetobacter calcoaceticus genes involved in biosynthesis of the coenzyme pyrri-  
 A/Reference number: A32252; MUID:89123056; PMID:2556663  
 A/Accession: A32252  
 A/Status: preliminary; not compared with conceptual translation  
 A/Residues: 1-204 <GOO>  
 A/Molecule type: DNA  
 A/Cross-references: UNIPROT:P07778; GB:X06452; NID:G38740; PIDN:CAA29752.1; PID:G38741  
 A/Experimental source: strain LMD 79.41  
 C/Comment: This sequence has motifs characteristic of a variety of phosphoesterases.  
 C/Superfamily: unassigned probable phosphoesterases; phosphoesterase core homology  
 C/Keywords: hydrolase  
 F:43-139/Domain: phosphoesterase core homology <PBC>

Query Match 70.5%; Score 31; DB 2; Length 204;  
 Best Local Similarity 62.5%; Pred. No. 41;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQLEPDAV 9  
 ::::|  
 Db 128 IQIDPDAV 135

RESULT 18  
 A40979  
 temperature shock-inducible protein T1P1 precursor - yeast (Saccharomyces cerevisiae)  
 N/Alternate names: protein YBR0622; protein YBR067c  
 C/Species: Saccharomyces cerevisiae  
 C/Date: 13-May-1992 #sequence\_revision 13-May-1992 #text\_change 09-Jul-2004  
 C/Accession: A40979; S45928; A57263  
 R/Kondo, K.; Inouye, M.  
 J. Biol. Chem. 266, 17537-17544, 1991  
 A>Title: T1P1, a cold shock-inducible gene of Saccharomyces cerevisiae.  
 A/Reference number: A40979; MUID:9137379; PMID:1894636  
 A/Accession: A40979  
 A/Molecule type: DNA  
 A/Residues: 1-210 <KON>  
 A/Cross-references: UNIPROT:P27654; GB:M71216; NID:G172983; PIDN:AAA35157.1; PID:G172984  
 R/Feldmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.  
 submitted to the Protein Sequence Database, August 1994  
 A/Reference number: S45927  
 A/Accession: S45928  
 A/Molecule type: DNA  
 A/Residues: 1-210 <PEL>  
 A/Cross-references: EMBL:Z35936; NID:G536313; PIDN:CAA85011.1; PID:G536314; MIPS:YBR067c  
 R/van Der Vaart, J.M.; Caro, L.H.P.; Chapman, J.W.; Kils, F.M.; Verrills, C.T.  
 J. Bacteriol. 177, 3104-3110, 1995  
 A>Title: Identification of three mannoproteins in the cell wall of Saccharomyces cerevisiae

A:Reference number: A57263; MUID:95286490; PMID:7768807  
A:Accession: A57263  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-210 <VAN>  
C:Genetics:  
A:Gene: SGD:TIPI  
A:Cross-references: SGD:S0000271; MIPS:YBR067C  
A:Map position: 2R  
C:Superfamily: serine-rich protein  
C:Keywords: cell wall; transmembrane protein  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-210/Product: temperature shock-inducible protein TIPI #status predicted <MAT>  
F:194-210/Domain: transmembrane #status predicted <TMM>

Query Match 70.5%; Score 31; DB 2; Length 210;  
Best Local Similarity 55.6%; Pred. No. 42;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDV 9  
|:|:|:|:  
|:|:|:|:  
Db 78 FSLBPDV 86

RESULT 19  
G70415  
nucleotide sugar epimerase - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C>Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C:Accession: G70415  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: G70415  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-321 <AQF>  
A:Cross-references: UNIPROT:O67354; GB:AE000735; NID:92983749; PIDN:AA07310.1; PID:9298  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: nse  
C:Superfamily: Escherichia coli UDPglucose 4-epimerase; UDPglucose 4-epimerase homology  
F:3-316/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 70.5%; Score 31; DB 2; Length 321;  
Best Local Similarity 66.7%; Pred. No. 68;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLOLEPDV 9  
|:|:|:|:  
|:|:|:|:  
Db 72 FOEFEPDV 80

RESULT 20  
T06786  
6a-hydroxymackin methyltransferase (EC 2.1.1.-) - garden pea  
C:Species: Pisum sativum (garden pea)  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T06786  
R:Wu, Q.; Preissig, C.L.; VanEtten, H.D.  
Plant Mol. Biol. 35, 551-560, 1997  
A:Title: Isolation of the cDNAs encoding (+)6a-hydroxymackin 3-O-methyltransferase.  
A:Reference number: Z15813; MUID:98009990; PMID:9349277  
A:Accession: T06786  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-360 <WU>  
A:Cross-references: UNIPROT:O24305; EMBL:U69554; NID:91568636; PIDN:AA049856.1; PID:9156  
A:Experimental source: cv. Alaska  
C:Genetics:  
A:Note: hmme

C:Function:  
A:Description: catalyzes the methylation of (+)6a-hydroxymackin to (+)piscatin  
A:Pathway: the terminal step in the piscatin biosynthesis  
C:Superfamily: O-methyltransferase  
C:Keywords: methyltransferase

Query Match 70.5%; Score 31; DB 2; Length 360;  
Best Local Similarity 75.0%; Pred. No. 77;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LOLEPDV 9  
|:|:|:|:  
|:|:|:|:  
Db 310 LOLEPDV 317

RESULT 21  
A86324  
protein F14D16.19 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: A86324  
R:Theologis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizer, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: A86324  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-384 <STO>  
A:Cross-references: UNIPROT:Q9LMK3; GB:AE005172; NID:98778280; PIDN:AA07289.1; GSPDB:GN  
C:Genetics:  
A:Gene: F14D16.19  
A:Map position: 1

Query Match 70.5%; Score 31; DB 2; Length 384;  
Best Local Similarity 71.4%; Pred. No. 83;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPD 7  
|:|:|:|:  
|:|:|:|:  
Db 91 FLELEPD 97

RESULT 22  
F81393  
probable alttronate hydrolase C-terminus Cj0483 [imported] - Campylobacter jejuni (strain  
C:Species: Campylobacter jejuni  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: F81393  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chillin  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Vanyilet, A.; Whitehead, S.; Barrel  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: F81393  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-390 <PAR>  
A:Cross-references: UNIPROT:Q9P126; GB:A113075; GB:A1111166; NID:96967817; PIDN:CAB7512  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: uxaA'; Cj0483

Query Match 70.5%; Score 31; DB 2; Length 390;  
Best Local Similarity 85.7%; Pred. No. 84;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 QLEFPAV 9  
|||  
Db 56 QLEFPAV 62

## RESULT 23

B90139  
Phosphomethylpyrimidine kinase (thid-1) [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C:Accession: B90139  
R:Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
A:Submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: B90139  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-400 <KUR>  
A:Cross-references: UNIPROT:Q981E3; GB:AE006641; NID:g13813129; PIDN:AAK40369.1; GSPDB:C  
A:Gene: thid-1

Query Match 70.5%; Score 31; DB 2; Length 400;  
Best Local Similarity 55.6%; Pred. No. 86;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLOEPAV 9  
|||  
Db 60 FLOEPAV 68

## RESULT 24

T4444  
hypothetical protein [imported] - Flavobacterium johnsoniae (fragment)  
C:Species: Flavobacterium johnsoniae  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T44444  
R:Agarwal, S.; Humicutt, D.W.; McBride, M.J.  
Proc. Natl. Acad. Sci. U.S.A. 94, 12139-12144, 1997  
A:Title: Cloning and characterization of the Flavobacterium johnsoniae (Cytophaga johns-  
on) reference number: 222770; MUID:98004537; PMID:9342376  
A:Accession: T44444  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-578 <AGA>  
A:Cross-references: UNIPROT:Q30490; EMBL:AF007381; NID:g2281662; PIDN:AAK19752.1; PID:g2  
A:Experimental source: strain ATCC17061

Query Match 70.5%; Score 31; DB 2; Length 578;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 QLEFPAV 9  
|||  
Db 278 QLEFPAV 284

## RESULT 25

T19678  
hypothetical protein C33D9.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T19678  
R:Lloyd, C.  
submitted to the EMBL Data Library, December 1995  
A:Reference number: Z19162  
A:Accession: T19678  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A:Residues: 1-875 <WIL>  
A:Cross-references: UNIPROT:Q18372; EMBL:Z68159; PIDN:CAA92282.1; GSPDB:GNO0022; CESP:C3  
A:Experimental source: clone C33D9  
C:Genetics:  
A:Gene: CESP:C33D9.1  
A:Map position: 4  
A:introns: 14/1; 74/3; 129/3; 171/3; 198/1; 241/2; 269/3; 352/3; 403/2; 515/3; 595/3; 65

Query Match 70.5%; Score 31; DB 2; Length 875;  
Best Local Similarity 85.7%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 QLEFPAV 9  
|||  
Db 47 QLEFPAV 53

## RESULT 26

G64514  
type I restriction enzyme homolog - Methanococcus jannaschii plasmid pURB800  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004  
C:Accession: G64514  
R:Bull, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Moese, C  
A:Title: Complete genome sequence of the mechanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: G64514  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1042 <BUL>  
A:Cross-references: UNIPROT:Q60295; GB:L77118; NID:g1500644; TIGR:MECL40; PIDN:AAK37109  
A:Genetics:  
A:Map position: ECLREVS2581-49453  
A:Genome: plasmid  
A:Note: this stable 58-kilobase pair plasmid is also designated ECL (large extrachromoso-  
me)  
C:Superfamily: type I site-specific deoxyribonuclease chain hsdR  
C:Keywords: ATP; nucleotide binding; P-loop  
F/336-343/Region: nucleotide-binding motif A (P-loop)  
F/439-440/Region: nucleotide-binding motif B  
F/439-442/Region: DEAH motif

Query Match 70.5%; Score 31; DB 2; Length 1042;  
Best Local Similarity 55.6%; Pred. No. 2.5e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gap 0;

OY 1 FLOEPAV 9  
|||  
Db 517 FLOEPAV 525

## RESULT 27

JC5546  
chitin synthase (BC 2.4.1.16) csma - Emerizella nidulans  
N:Alternate names: chitin-UDP N-acetylglucosaminyltransferase  
C:Species: Emerizella nidulans; Aspergillus nidulans  
C:Date: 02-Sep-1997 #sequence\_revision 05-Sep-1997 #text\_change 09-Jul-2004  
C:Accession: JC5546  
R:Fujiiwara, M.; Horiuchi, H.; Ohta, A.; Takagi, M.  
Biochem. Biophys. Res. Commun. 236, 75-78, 1997  
A:Title: A novel fungal gene encoding chitin synthase with a myosin motor-like domain.  
A:Reference number: JC5546; MUID:97366599; PMID:9223429  
A:Accession: JC5546  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-1852 <FUJ>  
A:Cross-references: UNIPROT:Q13281; DDBJ:AB000125; NID:g2308976; PID:dl022569; PID:g2308976  
C:Comment: This enzyme is a membrane-bound protein. It is involved in the catalytic polym-  
erization of N-acetylglucosamine.  
A:Gene: csma

C:Keywords: ATP, glycosyltransferase, hexosyltransferase  
F:102-110/Domain: ATP-binding #status predicted <ATP>

Query Match 70.5%; Score 31; DB 2; Length 1852;  
Best Local Similarity 62.5%; Pred. No. 4.8e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEPDA 8  
DB 158 FLELOVDA 165

## RESULT 28

hypothetical protein alr2139 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004

C:Accession: AE2073

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001

A>Title: Complete genomic sequence of the filamentous nitrogen-fixing Cyanobacterium Anabaena PCC 7120

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE2073

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-100 <KUR>

A:Cross-references: UNIPROT:Q8YV42; GB:BA000019; PIDD:BA073838.1; PID:g17131230; GSPDB:CA000001

A:Experimental source: strain PCC 7120

A:Gene: alr2139

Query Match 68.2%; Score 30; DB 2; Length 100;  
Best Local Similarity 85.7%; Pred. No. 30;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LOLEPDA 8  
DB 34 LRLEPDA 40

## RESULT 29

B84198

hypothetical protein Vmg0399h [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: B84198

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Letshauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabcic, Jung, K.H.; Alam, M.; Freitas, T. Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li, A. Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: B84198

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-155 <STO>

A:Cross-references: UNIPROT:Q9H555; GB:AE004437; NID:g10580011; PIDD:AA018953.1; GSPDB:CA000001

A:Gene: Vmg0399h

Query Match 68.2%; Score 30; DB 2; Length 155;  
Best Local Similarity 62.5%; Pred. No. 49;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LOLEPDAV 9  
DB 77 LQLEPDAI 84

## RESULT 30

D69888  
micrococcal nuclease homolog yncB - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C:Accession: D69888

R:Kumet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet, C.; Bron, S.; Broutillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chou, A.; Enllich, S.D.; Emerson, P.T.; Eutlan, K.D.; Ewington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gallert, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Y.M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portecelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schroeder, R.; Scofield, P.; Sekiguchi, J.; Sekowaka, A.; Serot, A.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumelein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: D69888

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-211 <KUN>

A:Cross-references: UNIPROT:P94492; GB:Z99113; GB:AL009126; NID:G2634090; PIDD:CA013646.1

A:Experimental source: strain 168

A:Gene: yncB

C:Superfamily: micrococcal nuclease

Query Match 68.2%; Score 30; DB 2; Length 211;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LOLEPDA 7  
DB 127 LQLEPDA 132

## RESULT 31

G45170

cytochrome-c oxidase (EC 1.9.3.1) chain II - ant (Lasius sp.) mitochondrion

C:Species: mitochondrion Lasius sp. (ant)

C:Date: 18-Jul-1995 #sequence\_revision 28-Jul-1995 #text\_change 18-Aug-2003

C:Accession: G45170

R:Li, H.; Beckenbach, A.T. Mol. Phylogenet. Evol. 1, 41-52, 1992

A:Title: Evolution of the mitochondrial cytochrome oxidase II gene among 10 orders of in

A:Reference number: A45170; MUID:94115687; PMID:1342923

A:Accession: G45170

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-226 <LIU>

A:Cross-references: GB:M83960; NID:g340659; PIDD:AAA1885.1; PID:g499645

A:Gene: COII

A:Genome: mitochondrion

C:Superfamily: cytochrome-c oxidase chain II, mitochondrial type; cytochrome-c oxidase c

C:Keywords: copper; electron transfer; membrane-associated complex; mitochondrial inner

F/8-213/Domain: cytochrome-c oxidase chain II homology <CO2>

F/160,195,199,206/Binding site: copper 1 (His, Cys, Cys, Met) #status predicted

F/195,197,199,203/Binding site: copper 2 (Cys, Glu, Cys, His) #status predicted

F/197/Binding site: magnesium (Glu) (shared with chain I) #status predicted

Query Match 68.2%; Score 30; DB 2; Length 226;  
Best Local Similarity 62.5%; Pred. No. 75;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLOLEPDA 8  
DB 112 FLNIFPDS 119

RESULT 32  
G90730  
antitermination protein [imported] - Escherichia coli (strain O157:H7, substrain RMD 05  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Nov-2003  
C:Accession: G90730  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: G90730  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-229 <HAV>  
A:Cross-references: GB:BA000007; PIDN:BA034238.1; PID:913360274; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: EC0815  
C:Superfamily: antitermination protein Q

Query Match 68.2%; Score 30; DB 2; Length 229;  
Best Local Similarity 66.7%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQEPDAV 9  
Db 198 FLQDSDAL 206

RESULT 33  
B81870  
probable membrane protein NMA1051 [imported] - Neisseria meningitidis (strain Z2491 sero  
C:Species: Neisseria meningitidis  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C:Accession: B81870  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jaseg, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A:Reference number: A81775; MUID:20222556; PMID:10761919  
A:Accession: B81870  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-237 <PAR>  
A:Cross-references: UNIPROT:Q9UV10; GB:AL162755; GB:AL157959; NID:97379742; PIDN:CAB8431  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: NMA1051

Query Match 68.2%; Score 30; DB 2; Length 237;  
Best Local Similarity 85.7%; Pred. No. 79;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QLEFDV 9  
Db 142 QLEFDV 148

RESULT 34  
D85581  
hypothetical protein Z0956 [imported] - Escherichia coli (strain O157:H7, substrain EDL5  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 03-Nov-2003  
C:Accession: D85581  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: D85581

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-257 <STO>  
A:Cross-references: GB:AE005174; NID:912513721; PIDN:AAG55112.1; GSPDB:GN00145; UWGP:Z09  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z0956  
C:Superfamily: antitermination protein Q

Query Match 68.2%; Score 30; DB 2; Length 257;  
Best Local Similarity 66.7%; Pred. No. 86;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLQEPDAV 9  
Db 226 FLQDSDAL 234

RESULT 35  
G82076  
3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17) cpda-type VC2433 [similarity] -  
N:Alternate names: icc protein  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: G82076  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;  
Hardison, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: G82076  
A:Molecule type: DNA  
A:Residues: 1-272 <HEI>  
A:Cross-references: UNIPROT:Q9KPD7; GB:AE004313; GB:AE003852; NID:9657007; PIDN:AAF957;  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC2433  
A:Map position: 1  
C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase, icc type; 3',5'-cyclic-nucleot  
C:Keywords: iron, metalloprotein, phosphoric diester hydrolase  
F13-202/Domain: 3',5'-cyclic-nucleotide phosphodiesterase cpda homology <CPDA>  
F13-93/Domain: phosphodiesterase core homology <PEC>

Query Match 68.2%; Score 30; DB 1; Length 272;  
Best Local Similarity 71.4%; Pred. No. 92;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QLEFDV 9  
Db 50 QLEFDV 56

RESULT 36  
T38816  
hypothetical protein SPAC4F10.12 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T38816  
R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, September 1997  
A:Reference number: Z21813  
A:Accession: T38816  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-280 <CON>  
A:Cross-references: UNIPROT:O36024; EMBL:Z96980; NID:e1060691; PIDN:CAB11715.1; GSPDB:GNC  
A:Experimental source: strain 972h-; cosmid c4F10  
C:Genetics:  
A:Gene: SPAC4F10.12  
A:Map position: 1

Query Match 68.2%; Score 30; DB 2; Length 280;

Best Local Similarity 85.7%; Pred. No. 95;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOEFPD 7  
Db 145 FLOEFPD 151

RESULT 37  
B97100  
pyridoxal kinase related protein [imported] - Clostridium acetobutylicum  
C/Species: Clostridium acetobutylicum  
C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C/Accession: B97100  
R/Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, O.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
A/Reference number: A96900; PMID:21359325; PMID:21359325  
A/Accession: B97100  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-290 <KUR>  
A/Cross-references: UNIPROT:Q97118; GB:AB001437; PIDN:AAK79589.1; PID:G15024579; GSPDB:G15024579  
A/Experimental source: Clostridium acetobutylicum ATCC824  
C/Genetics:  
A/Gene: CAC1622

Query Match 68.2%; Score 30; DB 2; Length 290;  
Best Local Similarity 62.5%; Pred. No. 99;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FLOEFPD 9  
Db 73 FLOEFPD 80

RESULT 38  
T12770  
probable endonuclease ynfK - Bacillus subtilis phage SPBc2  
C/Species: Bacillus subtilis phage SPBc2  
C/Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C/Accession: T12770; B69908  
R/Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Maue, C.; Karamata, D.  
submitted to the EMBL Data Library, August 1997  
A/Description: The complete nucleotide sequence of the Bacillus subtilis SPBc2 prophage  
A/Reference number: Z17583  
A/Accession: T12770  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-296 <LAZ>  
A/Cross-references: UNIPROT:Q64020; EMBL:AF020713; NID:G3025478; PID:G3025484; PIDN:AAK18Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Eutlian, K.D.; Errington, J.; Fabret, C.; Ferrati, E. Nature 390, 249-256, 1997  
A/Authors: Foulger, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.; Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue, Y.M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, T.; Winters, P.; Wipat, A.; Yamano, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K. A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A/Reference number: A69580; PMID:98044033; PMID:9384377  
A/Accession: B69908  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-296 <KUN>  
A/Cross-references: GB:Z99115; GB:AL009126; NID:G2634478; PIDN:CAB14079.1; PID:G2634581

A/Experimental source: strain 168  
C/Genetics:  
A/Gene: ynfK  
C/Superfamily: micrococcal nuclease

Query Match 68.2%; Score 30; DB 2; Length 296;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FLOEFPD 7  
Db 129 FLOEFPD 134

RESULT 39  
T23932  
hypothetical protein R05D7.4 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T23932  
R/Dobson, R.  
submitted to the EMBL Data Library, October 1996  
A/Reference number: Z19819  
A/Accession: T23932  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-299 <WIL>  
A/Cross-references: UNIPROT:Q45707; EMBL:Z81105; PIDN:CAB03219.1; GSPDB:GN00019; CESP:R05D7  
A/Experimental source: clone R05D7  
C/Genetics:  
A/Gene: CESP:R05D7.4  
A/Map position: 1  
A/Intons: 29/3; 122/1; 162/1; 231/2; 268/3  
C/Superfamily: tropinesterase

Query Match 68.2%; Score 30; DB 2; Length 299;  
Best Local Similarity 55.6%; Pred. No. 1e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLOEFPD 9  
Db 266 FPOVQPD 274

RESULT 40  
C69857  
formyltetrahydrofolate deformylase homolog ynfK - Bacillus subtilis  
C/Species: Bacillus subtilis  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Aug-2004  
C/Accession: C69857  
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Eutlian, K.D.; Errington, J.; Fabret, C.; Ferrati, E. Nature 390, 249-256, 1997  
A/Authors: Foulger, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.; Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue, Y.M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, T.; Winters, P.; Wipat, A.; Yamano, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K. A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A/Reference number: A69580; PMID:98044033; PMID:9384377  
A/Accession: C69857  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-300 <KUN>  
A/Cross-references: UNIPROT:Q34990; GB:Z99110; GB:AL009126; NID:G2633472; PIDN:CAB13168.  
A/Experimental source: strain 168  
C/Genetics:

A;Gene: ykkE  
 C;Superfamily: Formyltetrahydrofolate deformylase; phosphoribosylglycinamide formyltransferase  
 F;106-298/Domain: phosphoribosylglycinamide formyltransferase homology <PRGF>

Query March 68.2%; Score 30; DB 2; Length 300;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLQLEFD 7  
 ||::|||  
 Db 63 FLRIEFD 69

Search completed: January 12, 2005, 20:15:43  
 Job time : 18.4 secs

**THIS PAGE BLANK (USTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:38:08 ; Search time 85 Seconds  
(without alignments)  
60.922 Million cell updates/sec

Title: US-09-870-216C-5  
Perfect score: 44  
Sequence: 1 FLOLEPDAV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: UniProt\_02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_tramb1.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	84.1	124	094TV4	094tv4 haplogaeste
2	37	84.1	140	094TV9	094tv9 labidochiru
3	37	84.1	142	094TV8	094tv8 lopholithid
4	37	84.1	147	094TV6	094tv6 pagurus ber
5	37	84.1	155	094TV7	094tv7 cryptolitho
6	37	84.1	163	094TV6	094tv6 cryptolitho
7	37	84.1	166	094TV3	094tv3 lepidopa ca
8	37	84.1	170	094TV4	094tv4 blepharipod
9	37	84.1	171	094TV3	094tv3 paralomis g
10	37	84.1	172	094TV7	094tv7 oedignathus
11	37	84.1	174	094TV2	094tv2 phyllolitho
12	37	84.1	178	094TV3	094tv3 haplogaeste
13	37	84.1	179	094TV1	094tv1 rhinolithod
14	37	84.1	180	094TV2	094tv2 aegla scamo
15	37	84.1	180	094TV0	094tv0 lithodes sa
16	37	84.1	180	AA008408	Aa008408 aegla sca
17	37	84.1	181	094TV2	094tv2 lithodes ae
18	37	84.1	183	094TV5	094tv5 glyptolitho
19	37	84.1	183	094TV3	094tv3 raminoidea
20	37	84.1	184	094TV8	094tv8 munida quad
21	37	84.1	185	094TV7	094tv7 aegla ungu
22	37	84.1	188	094TV1	094tv1 lithodes ma
23	37	84.1	189	094TV5	094tv5 aegla sp. k
24	37	84.1	189	094TV0	094tv0 aegla sp. k
25	37	84.1	189	094TV3	094tv3 aegla neuqu
26	37	84.1	189	094TV1	094tv1 aegla inter
27	37	84.1	189	094TV4	094tv4 aegla inter
28	37	84.1	189	094TV5	094tv5 aegla inter
29	37	84.1	189	094TV6	094tv6 aegla plate
30	37	84.1	189	094TV7	094tv7 aegla plate
31	37	84.1	189	094TV9	094tv9 aegla humah

32	37	84.1	189	06PLU0	06plu0 aegla sp. k
33	37	84.1	189	06PLU1	06plu1 aegla sp. k
34	37	84.1	189	06PLU2	06plu2 aegla ringu
35	37	84.1	189	06PLU4	06plu4 aegla sanlo
36	37	84.1	189	06PLU5	06plu5 aegla sanlo
37	37	84.1	189	06PLU6	06plu6 aegla sanlo
38	37	84.1	189	06PLU8	06plu8 aegla septu
39	37	84.1	189	06PLU0	06plu0 aegla scamo
40	37	84.1	189	06PLV3	06plv3 aegla obacti
41	37	84.1	189	06PLV5	06plv5 aegla plate
42	37	84.1	189	06PLV6	06plv6 aegla plate
43	37	84.1	189	06PLV8	06plv8 aegla margi
44	37	84.1	189	06PLV9	06plv9 aegla strin
45	37	84.1	189	06PLW1	06plw1 aegla leptu

## ALIGNMENTS

RESULT 1					
ID	094TV4	PRELIMINARY	PRT:	124 AA.	
AC	094TV4				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
DE	Cytochrome oxidase subunit II (Fragment).				
OS	Haplogaster dentata.				
OC	Mitochondrion.				
OC	Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;				
OC	Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;				
OC	Lithodidae; Haplogaster.				
OX	NCBI_TaxID=174385;				
RP	[1]				
RP	SEQUENCE FROM N.A.				
RA	Zaklan S.D., Cunningham C.W.;				
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.				
CC	-1- FUNCTION: Cytochrome c oxidase is the component of the respiratory				
CC	chain that catalyzes the reduction of oxygen to water. Subunits 1-				
CC	3 form the functional core of the enzyme complex. Subunit 2				
CC	transfers the electrons from cytochrome c via its binuclear copper				
CC	A center to the binuclear center of the catalytic subunit 1 (By				
CC	similarity).				
CC	-1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome				
CC	c + 2 H(2)O.				
CC	-1- COPACTOR: Copper A (By similarity).				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial				
CC	inner membrane (By similarity).				
CC	-1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.				
DR	EMBL; AF425366; AAL26524.1; -.				
DR	GO; GO:0016020; C:membrane; IEA.				
DR	GO; GO:0005739; C:mitochondrion; IEA.				
DR	GO; GO:000507; F:copper ion binding; IEA.				
DR	GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.				
DR	GO; GO:0006118; P:electron transport; IEA.				
DR	InterPro: IPR008972; Cupredoxin.				
DR	InterPro: IPR003429; Cyt_c_ox_2.				
DR	Pfam; PF00116; COX2; 1.				
DR	Pfam; PF02790; COX2_TM; 1.				
DR	PRINTS; PR01166; CYCOXIDASEBII.				
KW	Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;				
KW	Oxidoreductase; Respiratory chain; Transmembrane; Transport.				
FT	NON_TER	1			
FT	NON_TER	1			
FT	NON_TER	1			
SO	SEQUENCE	124 AA;	14660 MW;	57DC335EF2212FB1 CRC64;	
Query Match					
Best Local Similarity 84.1%; Score 37; DB 2; Length 124;					
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;					
Qy	1 FLOLEPDA 8				
Db	93 FLOLEPDS 100				

## RESULT 2

094TU9 PRELIMINARY; PRT; 140 AA.  
 ID 094TU9;  
 AC 094TU9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Labyrinthus splendens.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
 OC Paguridae; Labyrinthus.  
 OX NCBI\_TaxID=174388;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zakian S.D., Cunningham C.W.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. Subunit 2 transfers the electrons from cytochrome c via its binuclear copper A center to the bimetallic center of the catalytic subunit 1 (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AF425371; AAL26529.1; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005507; F:copper ion binding; IEA.  
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001505; Copper Cua.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR InterPro; IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam; PF00116; COX2; 1.  
 DR PRINTS; PR01166; CYCOXIDASEII.  
 DR PRODOM; PD000131; Copper Cua; 1.  
 DR Kegg; Electron transport; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1 1  
 FT NON\_TER 140 140  
 SQ SEQUENCE 140 AA; 16429 MW; 92400CF6A7FEE00 CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 140;  
 Best Local Similarity 87.5%; Pred. No. 7; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 1;

Qy 1 FLOLEPDA 8  
 Db 97 FLOLEPDS 104

## RESULT 3

094TU8 PRELIMINARY; PRT; 142 AA.  
 ID 094TU8;  
 AC 094TU8;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Lopholittes manditi.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
 OC Lithodidae; Lopholittes.

OX NCBI\_TaxID=174396;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zakian S.D., Cunningham C.W.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. Subunit 2 transfers the electrons from cytochrome c via its binuclear copper A center to the bimetallic center of the catalytic subunit 1 (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AF425372; AAL26530.1; -;  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005507; F:copper ion binding; IEA.  
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001505; Copper Cua.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR InterPro; IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam; PF00116; COX2; 1.  
 DR PRINTS; PR01166; CYCOXIDASEII.  
 DR PRODOM; PD000131; Copper Cua; 1.  
 DR Kegg; Electron transport; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1 1  
 FT NON\_TER 142 142  
 SQ SEQUENCE 142 AA; 16485 MW; E00B56EB30703D02 CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 142;  
 Best Local Similarity 87.5%; Pred. No. 7.1; Mismatches 0; Indels 0; Gaps 0;  
 Matches 7; Conservative 1;

Qy 1 FLOLEPDA 8  
 Db 90 FLOLEPDS 97

## RESULT 4

094TU6 PRELIMINARY; PRT; 147 AA.  
 ID 094TU6;  
 AC 094TU6;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Pagurus bernhardus (common hermit crab).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
 OC Paguridae; Pagurus.  
 OX NCBI\_TaxID=174397;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zakian S.D., Cunningham C.W.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. Subunit 2 transfers the electrons from cytochrome c via its binuclear copper A center to the bimetallic center of the catalytic subunit 1 (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial

```

CC      inner membrane (By similarity).
CC      -1 SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR      EMBL; AF455374; AAL26532.1; -.
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0005739; C:mitochondrion; IEA.
DR      GO; GO:0005507; F:copper ion binding; IEA.
DR      GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR001505; Copper_CuA.
DR      InterPro; IPR002429; Cyt_c_ox_2.
DR      Pfam; PF00116; COX2; 1.
DR      Pfam; PF02790; COX2_TM; 1.
DR      PRINTS; PR01166; CYCOXIDASEII.
DR      ProDom; PD000131; Copper_CuA; 1.
KW      Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW      Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT      NON_TER
FT      NON_TER
SQ      SEQUENCE 147 AA; 17357 MW; 51BE16CED00DA004 CRC64;

Query Match      84.1%; Score 37; DB 2; Length 147;
Best Local Similarity 87.5%; Pred. No. 7.4;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLOLEPDA 8
      |||||:
      97 FLOLEPDS 104

RESULT 5
Q94TV7 PRELIMINARY; PRT; 155 AA.
ID Q94TV7
AC Q94TV7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Cryptolithodes sitchensis.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Cryptolithodes.
OX NCBI_TaxID=174327;
RN [1]
RP SEQUENCE FROM N.A.
RA Zaklan S.D., Cunningham C.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1 FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the binuclear center of the catalytic subunit 1 (By
CC similarity).
CC -1 CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1 COPFACTOR: Copper A (By similarity).
CC -1 SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1 SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR      EMBL; AF425363; AAL26521.1; -.
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0005739; C:mitochondrion; IEA.
DR      GO; GO:0005507; F:copper ion binding; IEA.
DR      GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR001505; Copper_CuA.
DR      InterPro; IPR002429; Cyt_c_ox_2.
DR      Pfam; PF00116; COX2; 1.
DR      Pfam; PF02790; COX2_TM; 1.
DR      PRINTS; PR01166; CYCOXIDASEII.
DR      ProDom; PD000131; Copper_CuA; 1.
KW      Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW      Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT      NON_TER
FT      NON_TER
SQ      SEQUENCE 163 AA; 18747 MW; B569F4CF5A80FDA8 CRC64;

Query Match      84.1%; Score 37; DB 2; Length 163;
Best Local Similarity 87.5%; Pred. No. 8.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLOLEPDA 8
      |||||:
      79 FLOLEPDS 86

```

```

KW      Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW      Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT      NON_TER
FT      NON_TER
SQ      SEQUENCE 155 AA; 17780 MW; 663801A37950E345 CRC64;

Query Match      84.1%; Score 37; DB 2; Length 155;
Best Local Similarity 87.5%; Pred. No. 7.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLOLEPDA 8
      |||||:
      76 FLOLEPDS 83

RESULT 6
Q94TV6 PRELIMINARY; PRT; 163 AA.
ID Q94TV6
AC Q94TV6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Cryptolithodes typicus.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Cryptolithodes.
OX NCBI_TaxID=174328;
RN [1]
RP SEQUENCE FROM N.A.
RA Zaklan S.D., Cunningham C.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1 FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the binuclear center of the catalytic subunit 1 (By
CC similarity).
CC -1 CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1 COPFACTOR: Copper A (By similarity).
CC -1 SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1 SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR      EMBL; AF425364; AAL26522.1; -.
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0005739; C:mitochondrion; IEA.
DR      GO; GO:0005507; F:copper ion binding; IEA.
DR      GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR001505; Copper_CuA.
DR      InterPro; IPR002429; Cyt_c_ox_2.
DR      InterPro; IPR002429; Cyt_c_ox_2.
DR      Pfam; PF00116; COX2; 1.
DR      Pfam; PF02790; COX2_TM; 1.
DR      PRINTS; PR01166; CYCOXIDASEII.
DR      ProDom; PD000131; Copper_CuA; 1.
KW      Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW      Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT      NON_TER
FT      NON_TER
SQ      SEQUENCE 163 AA; 18747 MW; B569F4CF5A80FDA8 CRC64;

Query Match      84.1%; Score 37; DB 2; Length 163;
Best Local Similarity 87.5%; Pred. No. 8.3;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FLOLEPDA 8
      |||||:
      79 FLOLEPDS 86

```

```

RESULT 7
08WFY3 PRELIMINARY; PRT; 166 AA.
AC 08WFY3;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Lepidoptera: California.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Hippoidea;
OC Alpheidae; Lepidoptera.
OX NCBI_TaxID=177216;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21884466; PubMed=11886621;
RA Morrison C.L.; Harvey A.W.; Lavery S.; Tieu K.; Huang Y.;
RA Cunningham C.W.;
RT "Mitochondrial gene rearrangements confirm the parallel evolution of
the crab-like form."
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:345-350(2002).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. Subunit 2
transfers the electrons from cytochrome c via its binuclear copper
A center to the bimetallic center of the catalytic subunit 1 (By
similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
EMBL; AF437625; AAL31588.1;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR PRODOM; PD000131; Copper_CuA; 1.
DR Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 166 166
FT SEQUENCE 166 AA; 19030 MW; C207194A21S9CDB2 CRC64;
SQ
Query Match 84.1%; Score 37; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 8.4;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FLOLPERDA 8
Db 89 FLOLPERDS 96

```

```

OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Hippoidea;
OC Alpheidae; Elephariopoda.
OX NCBI_TaxID=177216;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21884466; PubMed=11886621;
RA Morrison C.L.; Harvey A.W.; Lavery S.; Tieu K.; Huang Y.;
RA Cunningham C.W.;
RT "Mitochondrial gene rearrangements confirm the parallel evolution of
the crab-like form."
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:345-350(2002).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. Subunit 2
transfers the electrons from cytochrome c via its binuclear copper
A center to the bimetallic center of the catalytic subunit 1 (By
similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
EMBL; AF437625; AAL31588.1;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR PRODOM; PD000131; Copper_CuA; 1.
DR Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 170 170
FT SEQUENCE 170 AA; 19721 MW; D4F5895D1A87CA83 CRC64;
SQ
Query Match 84.1%; Score 37; DB 2; Length 170;
Best Local Similarity 87.5%; Pred. No. 8.6;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FLOLPERDA 8
Db 113 FLOLPERDS 120

```

```

CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL: AF425377; AAL26535.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR GO: GO:0004129; F:copper ion binding; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001505; Copper_CuA.
DR InterPro: IPR008972; Cupredoxin.
DR Pfam: PF00116; COX2; 1.
DR Pfam: PF02790; COX2_TM; 1.
DR PRINTS: PR01166; CYCOXIDASEII.
DR ProDom: PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT SEQUENCE 171 AA; 19619 MW; D26EBDDDD06235 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 171;
Best Local Similarity 87.5%; Pred. No. 8.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEFPA 8
Db 94 FLOLEFDS 101

RESULT 10
Q94TU7 PRELIMINARY; PRT; 172 AA.
AC Q94TU7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Oedignathus inermis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Oedignathus.
OX NCBI_TaxID=6743;
RN [1]
RP SEQUENCE FROM N.A.
RA Zaklan S.D., Cunningham C.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the binuclear center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL: AF425377; AAL26535.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR GO: GO:000507; F:copper ion binding; IEA.
DR GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001505; Copper_CuA.
DR InterPro: IPR008972; Cupredoxin.
DR Pfam: PF00116; COX2; 1.
DR Pfam: PF02790; COX2_TM; 1.
DR PRINTS: PR01166; CYCOXIDASEII.
DR ProDom: PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT SEQUENCE 174 AA; 19948 MW; EB262FBEF0CED389 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 174;
Best Local Similarity 87.5%; Pred. No. 8.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

DR InterPro: IPR002429; Cyt_c_ox_2.
DR Pfam: PF00116; COX2; 1.
DR Pfam: PF02790; COX2_TM; 1.
DR PRINTS: PR01166; CYCOXIDASEII.
DR ProDom: PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT SEQUENCE 172 AA; 19918 MW; BA397959C0867049 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 172;
Best Local Similarity 87.5%; Pred. No. 8.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEFPA 8
Db 97 FLOLEFDS 104

RESULT 11
Q94TU2 PRELIMINARY; PRT; 174 AA.
AC Q94TU2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
OS Phyllolithodes papillosus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Phyllolithodes.
OX NCBI_TaxID=17407;
RN [1]
RP SEQUENCE FROM N.A.
RA Zaklan S.D., Cunningham C.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the binuclear center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL: AF425378; AAL26536.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR GO: GO:000507; F:copper ion binding; IEA.
DR GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001505; Copper_CuA.
DR InterPro: IPR008972; Cupredoxin.
DR Pfam: PF00116; COX2; 1.
DR Pfam: PF02790; COX2_TM; 1.
DR PRINTS: PR01166; CYCOXIDASEII.
DR ProDom: PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT SEQUENCE 174 AA; 19948 MW; EB262FBEF0CED389 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 174;
Best Local Similarity 87.5%; Pred. No. 8.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 FLOLEPDA 8  
 |||||:  
 Db 90 FLOLEPDS 97

## RESULT 12

Q94TV3 PRELIMINARY; PRT; 178 AA.  
 ID 094TV3  
 AC 094TV3  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Haplogaster mertensii.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
 OC Lithodidae; Haplogaster.  
 OX NCBI\_TaxID=174386;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zaklan S.D., Cunningham C.W.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 3 form the functional core of the enzyme complex. Subunit 2  
 transfers the electrons from cytochrome c via its binuclear copper  
 A center to the bimetallic center of the catalytic subunit 1 (By  
 similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AF425367; AAL26525.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005507; F:copper ion binding; IEA.  
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001505; Copper\_CuA.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PR01166; CYCOXIDASEIT.  
 DR ProDom; PD000131; Copper\_CuA; 1.  
 DR Kegg; K00001; Electron transport; Inner membrane; Mitochondrion;  
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1 1  
 FT NON\_TER 178 178  
 SQ SEQUENCE 178 AA; 20305 MW; D21D40F3CE64D94F CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 178;  
 Best Local Similarity 87.5%; Pred. No. 9.1;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

## RESULT 13

Q94TV1 PRELIMINARY; PRT; 179 AA.  
 ID 094TV1  
 AC 094TV1  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Rhinolithodes wosnesenskii.  
 OG Mitochondrion.

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
 OC Lithodidae; Rhinolithodes.  
 OX NCBI\_TaxID=174409;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zaklan S.D., Cunningham C.W.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 3 form the functional core of the enzyme complex. Subunit 2  
 transfers the electrons from cytochrome c via its binuclear copper  
 A center to the bimetallic center of the catalytic subunit 1 (By  
 similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AF425379; AAL26537.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005507; F:copper ion binding; IEA.  
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001505; Copper\_CuA.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PR01166; CYCOXIDASEIT.  
 DR ProDom; PD000131; Copper\_CuA; 1.  
 DR Kegg; K00001; Electron transport; Inner membrane; Mitochondrion;  
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1 1  
 FT NON\_TER 179 179  
 SQ SEQUENCE 179 AA; 20598 MW; 18058F408B1B2A8 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 179;  
 Best Local Similarity 87.5%; Pred. No. 9.1;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEPDA 8  
 |||||:  
 Db 93 FLOLEPDS 100

## RESULT 14

Q6PLV2 PRELIMINARY; PRT; 180 AA.  
 ID 06PLV2  
 AC 06PLV2  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 GN Name=COII;  
 OS Aegla scamosa.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 OC Galatheoidea; Aegilidae; Aegla.  
 OX NCBI\_TaxID=273843;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;  
 RT "Molecular systematics and biogeography of the southern South American  
 freshwater crabs Aegla (Decapoda: Anomura: Aegilidae) using multiple  
 heuristic tree search approaches.";  
 RL Syst. Biol. 53:0-0(2004).  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunit 2  
 3 form the functional core of the enzyme complex. Subunit 2

transfers the electrons from cytochrome c via its binuclear copper A center to the bimetallic center of the catalytic subunit 1 (By similarity).  
 -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.  
 -1- COPACTOR: Copper A (By similarity).  
 -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).  
 -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 EMBL: AY595772; AAT08408.1; -.  
 DR GO: GO:0005739; C:mitochondrion; IEA.  
 DR InterPro: IPR001505; Copper\_CuA.  
 DR InterPro: IPR008972; Cupredoxin.  
 DR Pfam: PF00116; COX2; 1.  
 DR Pfam: PF02790; COX2\_TM; 1.  
 DR PRINTS: PR01166; CYCOXIDASEII.  
 DR ProDom: PD000131; Copper\_CuA; 1.  
 DR Copper: Electron transport; Inner membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1 180  
 FT NON\_TER 1 180  
 SQ SEQUENCE 180 AA; 20778 MW; C627D7E7F4CA2C0B CRC64;

Query Match 84.1%; Score 37; DB 2; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 9.2;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEPDA 8  
 Db 92 FLOLEPDS 99

## RESULT 15

094TV0 PRELIMINARY; PRT; 180 AA.  
 AC 094TV0;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Lichodes santolla.  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
 OC Lichodidae; Lichodes.  
 OC NCBI\_TaxID=174394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zaklan S.D., Cunningham C.W.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. Subunit 2 transfers the electrons from cytochrome c via its binuclear copper A center to the bimetallic center of the catalytic subunit 1 (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.  
 CC -1- COPACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 EMBL: AF425370; AAL26528.1; -.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0005739; C:mitochondrion; IEA.  
 DR GO: GO:0005507; F:copper ion binding; IEA.  
 DR GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.  
 DR GO: GO:0006118; P:electron transport; IEA.  
 DR InterPro: IPR001505; Copper\_CuA.  
 DR InterPro: IPR008972; Cupredoxin.  
 DR InterPro: IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam: PF00116; COX2; 1.

DR Pfam: PF02790; COX2\_TM; 1.  
 DR PRINTS: PR01166; CYCOXIDASEII.  
 DR ProDom: PD000131; Copper\_CuA; 1.  
 KW Copper: Electron transport; Inner membrane; Membrane; Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1 180  
 FT NON\_TER 1 180  
 SQ SEQUENCE 180 AA; 20816 MW; FAED7795D4F090F1 CRC64;

Qy 1 FLOLEPDA 8  
 Db 97 FLOLEPDS 104

## RESULT 16

AAT08408 PRELIMINARY; PRT; 180 AA.  
 AC AAT08408;  
 DT 12-MAY-2004 (TREMBLrel. 27, Created)  
 DT 12-MAY-2004 (TREMBLrel. 27, Last sequence update)  
 DT 12-MAY-2004 (TREMBLrel. 27, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 GN COIL.  
 OS Aegla scamosa.  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 OC Galatheoidea; Aegidae; Aegla.  
 OC NCBI\_TaxID=273843;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Perez-Iosada M., Bond-Buckup G., Jara C.G., Crandall K.A.;  
 RT "Molecular systematics and biogeography of the southern South American freshwater crabs Aegla (Decapoda: Anomura: Aegidae) using multiple heuristic tree search approaches."  
 RL Syst. Biol. 53:0-0(2004).  
 DR EMBL: AY595772; AAT08408.1; -.  
 DR Mitochondrion.  
 FT NON\_TER 1 180  
 FT NON\_TER 1 180  
 SQ SEQUENCE 180 AA; 20778 MW; C627D7E7F4CA2C0B CRC64;

Query Match 84.1%; Score 37; DB 2; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 9.2;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEPDA 8  
 Db 92 FLOLEPDS 99

## RESULT 17

094TV2 PRELIMINARY; PRT; 181 AA.  
 AC 094TV2;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Lichodes aequispinus.  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;  
 OC Lichodidae; Lichodes.  
 OC NCBI\_TaxID=174392;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Zaklan S.D., Cunningham C.W.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

```

CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL: AF425368; AAL26526.1; -.
CC GO: GO:0016020; C:membrane; IEA.
CC GO: GO:0005739; C:mitochondrion; IEA.
CC GO: GO:0005507; F:copper ion binding; IEA.
CC GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
CC GO: GO:0006118; P:electron transport; IEA.
CC InterPro: IPR001505; Copper_CuA.
CC InterPro: IPR008972; Cupredoxin.
CC InterPro: IPR002429; Cyt_c_ox_2.
CC Pfam: PF00116; COX2; 1.
CC Pfam: PF02790; COX2_TM; 1.
CC PRINTS: PR01166; CYCOXIDASEII.
CC ProDom: PD000131; Copper_CuA; 1.
CC Copper; Electron transport; Inner membrane; Mitochondrion;
CC Oxidoreductase; Respiratory chain; Transmembrane; Transport.
CC NON_TER 1 1
CC NON_TER 181 181
CC SEQUENCE 181 AA; 20842 MW; 0FBE984FAAF90D8 CRC64;
CC
CC Query Match 84.1%; Score 37; DB 2; Length 181;
CC Best Local Similarity 87.5%; Pred. No. 9.2;
CC Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 FLOLEFPA 8
CC DB 95 FLOLEFDS 102
CC
CC RESULT 18
CC Q94TV5 PRELIMINARY; PRT; 183 AA.
CC ID Q94TV5;
CC DT 01-BEC-2001 (T-EMBLrel. 19, Created)
CC DT 01-BEC-2001 (T-EMBLrel. 19, Last sequence update)
CC DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
CC DE Cytochrome oxidase subunit II (Fragment).
CC OS Glyptolithodes cristatipes.
CC Mitochondrion.
CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
CC Lithodidae; Glyptolithodes.
CC NCBI_TaxID=174383;
CC
CC SEQUENCE FROM N.A.
CC Zaklan S.D.; Cunningham C.W.;
CC Subtilan S.D.; EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL: AF425365; AAL26523.1; -.
CC GO: GO:0016020; C:membrane; IEA.
CC GO: GO:0005739; C:mitochondrion; IEA.

```

```

DR GO: GO:0005507; F:copper ion binding; IEA.
DR GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001505; Copper_CuA.
DR InterPro: IPR008972; Cupredoxin.
DR InterPro: IPR002429; Cyt_c_ox_2.
DR Pfam: PF00116; COX2; 1.
DR Pfam: PF02790; COX2_TM; 1.
DR PRINTS: PR01166; CYCOXIDASEII.
DR ProDom: PD000131; Copper_CuA; 1.
CC Copper; Electron transport; Inner membrane; Mitochondrion;
CC Oxidoreductase; Respiratory chain; Transmembrane; Transport.
CC NON_TER 1 1
CC NON_TER 183 183
CC SEQUENCE 183 AA; 21102 MW; C15C303F66EDE4F0 CRC64;
CC
CC Query Match 84.1%; Score 37; DB 2; Length 183;
CC Best Local Similarity 87.5%; Pred. No. 9.3;
CC Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 FLOLEFPA 8
CC DB 97 FLOLEFDS 104
CC
CC RESULT 19
CC Q9MEZ3 PRELIMINARY; PRT; 183 AA.
CC ID Q9MEZ3;
CC DT 01-MAR-2002 (T-EMBLrel. 20, Created)
CC DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
CC DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
CC DE Cytochrome oxidase subunit II (Fragment).
CC OS Ranioides louisianensis.
CC Mitochondrion.
CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
CC Eubrachyura; Ranioides; Raniidae; Ranioides.
CC NCBI_TaxID=6794;
CC
CC SEQUENCE FROM N.A.
CC MEDLINE=21884466; PubMed=11886621;
CC Morrison C.L.; Harvey A.W.; Lavery S.; Tieu K.; Huang Y.;
CC Cunningham C.W.;
CC "Mitochondrial gene rearrangements confirm the parallel evolution of
CC the crab-like form."
CC Proc. R. Soc. Lond., B. Biol. Sci. 269:345-350(2002).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL: AF437617; AAL31579.1; -.
CC GO: GO:0016020; C:membrane; IEA.
CC GO: GO:0005739; C:mitochondrion; IEA.
CC GO: GO:0005507; F:copper ion binding; IEA.
CC GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
CC GO: GO:0006118; P:electron transport; IEA.
CC InterPro: IPR001505; Copper_CuA.
CC InterPro: IPR008972; Cupredoxin.
CC InterPro: IPR002429; Cyt_c_ox_2.
CC Pfam: PF00116; COX2; 1.
CC Pfam: PF02790; COX2_TM; 1.
CC PRINTS: PR01166; CYCOXIDASEII.
CC ProDom: PD000131; Copper_CuA; 1.
CC Copper; Electron transport; Inner membrane; Mitochondrion;

```



KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON TER 1  
 FT NON TER 183  
 SQ SEQUENCE 183 AA; 21190 MW; 083F7ACBA63B6518 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 183;  
 Best Local Similarity 87.5%; Pred. No. 9.3;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEFPA 8  
 |||||  
 DB 99 FLOLEFDS 106

## RESULT 20

O8WFY8 PRELIMINARY; PRT; 184 AA.

AC O8WFY8 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Munda quadripinna.

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 OC Galatheoidea; Galatheidae; Munda.  
 OX NCBI\_TaxID=177235;

RN NCI  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21884466; PubMed=11886621;  
 RA Morrison C.L., Harvey A.W., Lavery S., Tieu K., Huang Y.,  
 RA Cunningham C.W.;

RT "Mitochondrial gene rearrangements confirm the parallel evolution of  
 the crab-like form.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:345-350(2002).

CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 3 form the functional core of the enzyme complex. Subunit 2  
 transfers the electrons from cytochrome c via its binuclear copper  
 A center to the binuclear center of the catalytic subunit 1 (By  
 similarity).

CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 c + 2 H(2)O.

CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 inner membrane (By similarity).

CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AF437622; AAL31584.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005507; F:copper ion binding; IEA.  
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001505; Copper\_CuA.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR InterPro; IPR002429; Cyt\_c\_ox\_2.

DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PRO1166; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CuA; 1.  
 KW Copper; Electron transport; Inner membrane; Mitochondrion;  
 KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.

FT NON TER 1  
 FT NON TER 184  
 SQ SEQUENCE 184 AA; 21168 MW; 45AD6A9308516F1B CRC64;

Query Match 84.1%; Score 37; DB 2; Length 184;  
 Best Local Similarity 87.5%; Pred. No. 9.4;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEFPA 8  
 |||||

DB 99 FLOLEFDS 106

## RESULT 21

O8WFY7 PRELIMINARY; PRT; 185 AA.

AC O8WFY7 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 OS Aegla uruguayana.

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 OC Galatheoidea; Aegllidae; Aegla.  
 OX NCBI\_TaxID=177212;

RN NCI  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21884466; PubMed=11886621;  
 RA Morrison C.L., Harvey A.W., Lavery S., Tieu K., Huang Y.,  
 RA Cunningham C.W.;

RT "Mitochondrial gene rearrangements confirm the parallel evolution of  
 the crab-like form.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:345-350(2002).

CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 3 form the functional core of the enzyme complex. Subunit 2  
 transfers the electrons from cytochrome c via its binuclear copper  
 A center to the binuclear center of the catalytic subunit 1 (By  
 similarity).

CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 c + 2 H(2)O.

CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 inner membrane (By similarity).

CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AF437623; AAL31585.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR GO; GO:0005507; F:copper ion binding; IEA.  
 DR GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001505; Copper\_CuA.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR InterPro; IPR002429; Cyt\_c\_ox\_2.

DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PRO1166; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CuA; 1.

KW Copper; Electron transport; Inner membrane; Mitochondrion;  
 KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.

FT NON TER 1  
 FT NON TER 185  
 SQ SEQUENCE 185 AA; 21419 MW; 6A0C538A7D8A6F82 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 185;  
 Best Local Similarity 87.5%; Pred. No. 9.5;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEFPA 8  
 |||||  
 DB 100 FLOLEFDS 107

## RESULT 22

O94TV1 PRELIMINARY; PRT; 188 AA.

AC O94TV1 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).

```

OS Lithodes maja.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura; Paguroidea;
OC Lithodidae; Lithodes.
OX NCBI_TaxID=174393;
RN [1]
RP SEQUENCE FROM N.A.
RA Zakian S.D., Cunningham C.W.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL: AF425369; AAL26527.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR GO: GO:0005507; F:copper ion binding; IEA.
DR GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001505; Copper_CuA.
DR InterPro: IPR008972; Cupredoxin.
DR Pfam: PF00116; COX2; 1.
DR Pfam: PF02790; COX2 TM; 1.
DR PRINTS: PR01166; CYCOXIDASEIT.
DR ProDom: PD000131; Copper_CuA; 1.
DR Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 188
FT SEQUENCE 188 AA; 21638 MW; 8F6D57A18ECCAE5C CRC64;
SQ
Query Match 84.1%; Score 37; DB 2; Length 188;
Best Local Similarity 87.5%; Pred. No. 9.6;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FLOLEPDA 8
DB 102 FLOLEPDS 109
RESULT 23
Q6PLS9 PRELIMINARY; PRT; 189 AA.
ID Q6PLS9;
AC Q6PLS9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Aegla sp. KACA0541.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;
OC Galatheoidea; Aegllidae; Aegla.
OX NCBI_TaxID=274482;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RT "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegllidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
DE -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory

```

```

CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL: AY595795; AAT08431.1; -.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR InterPro: IPR001505; Copper_CuA.
DR InterPro: IPR008972; Cupredoxin.
DR InterPro: IPR002429; Cyt_c_ox_2.
DR Pfam: PF00116; COX2; 1.
DR Pfam: PF02790; COX2 TM; 1.
DR PRINTS: PR01166; CYCOXIDASEIT.
DR ProDom: PD000131; Copper_CuA; 1.
DR Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 189
FT SEQUENCE 189 AA; 21586 MW; 48B356FEDBDEBFCAD CRC64;
SQ
Query Match 84.1%; Score 37; DB 2; Length 189;
Best Local Similarity 87.5%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FLOLEPDA 8
DB 101 FLOLEPDS 108
RESULT 24
Q6PLT0 PRELIMINARY; PRT; 189 AA.
ID Q6PLT0;
AC Q6PLT0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Aegla sp. KACA0538.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;
OC Galatheoidea; Aegllidae; Aegla.
OX NCBI_TaxID=274481;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RT "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegllidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
DE -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL: AY595794; AAT08430.1; -.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR InterPro: IPR001505; Copper_CuA.
DR InterPro: IPR008972; Cupredoxin.

```

```

DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR ProDom; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 189 AA; 21567 MW; 5D5356FEE7F1F6B7 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 189;
Best Local Similarity 87.5%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEFDA 8
Db 101 FLOLEFDS 108

RESULT 25
Q6PLT1 PRELIMINARY; PRT; 189 AA.
AC Q6PLT1;
DT 05-JUL-2004 (TRENBLREL. 27, Created)
DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Aegla neuquensis.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;
OC Galatheoidea; Aegliidae; Aegla.
OX NCBI_TaxID=273850;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RT "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL; AY55793; AAT08428.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR ProDom; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 189 AA; 21749 MW; F740F0F77889210D CRC64;

Query Match 84.1%; Score 37; DB 2; Length 189;
Best Local Similarity 87.5%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 FLOLEFDA 8
Db 101 FLOLEFDS 108

RESULT 26
Q6PLT3 PRELIMINARY; PRT; 189 AA.
AC Q6PLT3;
DT 05-JUL-2004 (TRENBLREL. 27, Created)
DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Aegla intercalata.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;
OC Galatheoidea; Aegliidae; Aegla.
OX NCBI_TaxID=273849;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RT "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL; AY55793; AAT08427.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR ProDom; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 189 AA; 21731 MW; 76055B4098673CE0 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 189;
Best Local Similarity 87.5%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEFDA 8
Db 101 FLOLEFDS 108

RESULT 27
Q6PLT4 PRELIMINARY; PRT; 189 AA.
AC Q6PLT4;
DT 05-JUL-2004 (TRENBLREL. 27, Created)
DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;

```

```

OS Aegla intercalata.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;
OC Galatheoidea; Aegliidae; Aegla.
OX NCBI_TaxID=273849;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RT "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC DR EMBL; AY595789; AAT08426.1; -.
CC DR GO; GO:0005739; C:mitochondrion; IEA.
CC DR InterPro; IPR001505; Copper_CuA.
CC DR InterPro; IPR008972; Cupredoxin.
CC DR Pfam; PF00116; COX2; 1.
CC DR Pfam; PF02790; COX2_TM; 1.
CC DR PRINTS; PR01166; CYCOXIDASEII.
CC DR ProDom; PD000131; Copper_CuA; 1.
CC DR Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
CC KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.
CC FT NON TER 1 189
CC FT SEQUENCE 189 AA; 21745 MW; 11C1F5D5CADA40CE CRC64;
SQ
Query Match 84.1%; Score 37; DB 2; Length 189;
Best Local Similarity 87.5%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 F1QLEFDA 8
Db 101 F1QLEFDS 108
RESULT 28
O6PLT5 PRELIMINARY; PRT; 189 AA.
ID O6PLT5;
AC O6PLT5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Aegla intercalata.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;
OC Galatheoidea; Aegliidae; Aegla.
OX NCBI_TaxID=273849;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RT "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC DR EMBL; AY595788; AAT08424.1; -.
CC DR GO; GO:0005739; C:mitochondrion; IEA.
CC DR InterPro; IPR001505; Copper_CuA.
CC DR InterPro; IPR008972; Cupredoxin.
CC DR InterPro; IPR002429; Cyt_c_ox_2.

```

```

CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL; AY595789; AAT08425.1; -.
CC DR GO; GO:0005739; C:mitochondrion; IEA.
CC DR InterPro; IPR001505; Copper_CuA.
CC DR InterPro; IPR008972; Cupredoxin.
CC DR Pfam; PF00116; COX2; 1.
CC DR Pfam; PF02790; COX2_TM; 1.
CC DR PRINTS; PR01166; CYCOXIDASEII.
CC DR ProDom; PD000131; Copper_CuA; 1.
CC DR Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
CC KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.
CC FT NON TER 1 189
CC FT SEQUENCE 189 AA; 21749 MW; 4E3D31809CB54CC CRC64;
SQ
Query Match 84.1%; Score 37; DB 2; Length 189;
Best Local Similarity 87.5%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 F1QLEFDA 8
Db 101 F1QLEFDS 108
RESULT 29
O6PLT6 PRELIMINARY; PRT; 189 AA.
ID O6PLT6;
AC O6PLT6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Aegla platensis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;
OC Galatheoidea; Aegliidae; Aegla.
OX NCBI_TaxID=273811;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RT "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the bimetallic center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
CC EMBL; AY595788; AAT08424.1; -.
CC DR GO; GO:0005739; C:mitochondrion; IEA.
CC DR InterPro; IPR001505; Copper_CuA.
CC DR InterPro; IPR008972; Cupredoxin.
CC DR InterPro; IPR002429; Cyt_c_ox_2.

```

```

DR PFam; PF00116; COX2.1.
DR PFam; PF02790; COX2.TM; 1.
DR PRINTS; PRO1166; CYCOXIDASEII.
DR ProDom; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 189
SQ SEQUENCE 189 AA; 21791 MW; 8F22AF85D97325BE CRC64;

Query Match
Best Local Similarity 84.1%; Score 37; DB 2; Length 189;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEPDA 8
Db 101 FLOLEPDS 108

RESULT 30
O6PLT7 PRELIMINARY; PRT; 189 AA.
AC O6PLT7;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Aegla platensis.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Decapoda; Decapoda; Pleocyemata; Anomura;
OC Galatheidae; Aegliidae; Aegla.
OX NCBI_TaxID=273811;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Loesada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RT "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the binuclear center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COPFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL; AY595787; AAT08423.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR InterPro; IPR01505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF00116; COX2.1.
DR Pfam; PF02790; COX2.TM; 1.
DR PRINTS; PRO1166; CYCOXIDASEII.
DR ProDom; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 189
SQ SEQUENCE 189 AA; 21807 MW; 8F34758B636B45BE CRC64;

Query Match
Best Local Similarity 84.1%; Score 37; DB 2; Length 189;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEPDA 8
Db 101 FLOLEPDS 108

```

```

Db 101 FLOLEPDS 108
OY 1 FLOLEPDA 8
Db 101 FLOLEPDS 108

RESULT 31
O6PLT9 PRELIMINARY; PRT; 189 AA.
AC O6PLT9;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Aegla humahuaca.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Decapoda; Decapoda; Pleocyemata; Anomura;
OC Galatheidae; Aegliidae; Aegla.
OX NCBI_TaxID=273848;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Loesada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RT "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the binuclear center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COPFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL; AY595785; AAT08421.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR InterPro; IPR01505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR002429; Cyt_c_ox_2.
DR Pfam; PF00116; COX2.1.
DR Pfam; PF02790; COX2.TM; 1.
DR PRINTS; PRO1166; CYCOXIDASEII.
DR ProDom; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 189
SQ SEQUENCE 189 AA; 21731 MW; 76055B4098673CE0 CRC64;

Query Match
Best Local Similarity 84.1%; Score 37; DB 2; Length 189;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLOLEPDA 8
Db 101 FLOLEPDS 108

RESULT 32
O6PLU0 PRELIMINARY; PRT; 189 AA.
AC O6PLU0;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Aegla sp. KACa0480.

```

OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 CC Galatheoidea; Aegidae; Aegla.  
 OX NCBI\_TaxID=274480;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;  
 RT "Molecular systematics and biogeography of the southern South American  
 freshwater crabs Aegla (Decapoda: Anomura: Aegidae) using multiple  
 RT heuristic tree search approaches."  
 RL Syst. Biol. 53:0-0(2004).  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 CC 3 form the functional core of the enzyme complex. Subunit 2  
 CC transfers the electrons from cytochrome c via its binuclear copper  
 CC A center to the bimetallic center of the catalytic subunit 1 (By  
 CC similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 CC c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 CC EMBL: AY595784; AA08420.1; -.  
 CC GO: GO:0005739; C:mitochondrion; IEA.  
 CC InterPro: IPR001505; Copper Cua.  
 CC InterPro: IPR008972; Cupredoxin.  
 CC Pfam: PF02790; COX2; 1.  
 CC Cyt\_c\_ox\_2.  
 DR InterPro: IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam: PF02790; COX2; 1.  
 DR PRINTS: PR01166; CYCOXIDASEII.  
 DR PRODOM: PD000131; Copper Cua; 1.  
 KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 189 189  
 FT NON\_TER 1 1  
 SQ SEQUENCE 189 AA; 21685 MW; C737BD06415DB681 CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 189;  
 Best Local Similarity 87.5%; Pred. No. 9.7;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 FLOLFDA 8  
 DB 101 FLOLFDS 108  
 RESULT 33  
 O6PLU1 PRELIMINARY; PRT; 189 AA.  
 ID O6PLU1;  
 AC O6PLU1;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 GN Name=COII;  
 OS Aegla sp. KACa0479.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 CC Galatheoidea; Aegidae; Aegla.  
 OX NCBI\_TaxID=274479;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;  
 RT "Molecular systematics and biogeography of the southern South American  
 freshwater crabs Aegla (Decapoda: Anomura: Aegidae) using multiple  
 RT heuristic tree search approaches."  
 RL Syst. Biol. 53:0-0(2004).  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 CC 3 form the functional core of the enzyme complex. Subunit 2  
 CC transfers the electrons from cytochrome c via its binuclear copper  
 CC A center to the bimetallic center of the catalytic subunit 1 (By  
 CC similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 CC c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 CC EMBL: AY595782; AA08418.1; -.  
 CC GO: GO:0005739; C:mitochondrion; IEA.  
 CC InterPro: IPR001505; Copper Cua.  
 CC InterPro: IPR008972; Cupredoxin.  
 CC InterPro: IPR002429; Cyt\_c\_ox\_2.

CC transfers the electrons from cytochrome c via its binuclear copper  
 CC A center to the bimetallic center of the catalytic subunit 1 (By  
 CC similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 CC c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 CC EMBL: AY595783; AA08419.1; -.  
 CC GO: GO:0005739; C:mitochondrion; IEA.  
 CC InterPro: IPR001505; Copper Cua.  
 CC InterPro: IPR008972; Cupredoxin.  
 CC Pfam: PF02790; COX2; 1.  
 CC Cyt\_c\_ox\_2.  
 DR InterPro: IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam: PF02790; COX2; 1.  
 DR PRINTS: PR01166; CYCOXIDASEII.  
 DR PRODOM: PD000131; Copper Cua; 1.  
 KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 189 189  
 FT NON\_TER 1 1  
 SQ SEQUENCE 189 AA; 21685 MW; C737BD06415DB681 CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 189;  
 Best Local Similarity 87.5%; Pred. No. 9.7;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 FLOLFDA 8  
 DB 101 FLOLFDS 108  
 RESULT 34  
 O6PLU2 PRELIMINARY; PRT; 189 AA.  
 ID O6PLU2;  
 AC O6PLU2;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 GN Name=COII;  
 OS Aegla ringueleti.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 CC Galatheoidea; Aegidae; Aegla.  
 OX NCBI\_TaxID=273847;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;  
 RT "Molecular systematics and biogeography of the southern South American  
 freshwater crabs Aegla (Decapoda: Anomura: Aegidae) using multiple  
 RT heuristic tree search approaches."  
 RL Syst. Biol. 53:0-0(2004).  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 CC 3 form the functional core of the enzyme complex. Subunit 2  
 CC transfers the electrons from cytochrome c via its binuclear copper  
 CC A center to the bimetallic center of the catalytic subunit 1 (By  
 CC similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 CC c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 CC EMBL: AY595782; AA08418.1; -.  
 CC GO: GO:0005739; C:mitochondrion; IEA.  
 CC InterPro: IPR001505; Copper Cua.  
 CC InterPro: IPR008972; Cupredoxin.  
 CC InterPro: IPR002429; Cyt\_c\_ox\_2.

DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PR0116; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CUA; 1.  
 DR Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1 1  
 FT NON\_TER 189 189  
 SQ SEQUENCE 189 AA; 21715 MW; F83745670E3A5C19 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 189;  
 Best Local Similarity 87.5%; Pred. No. 9.7;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEPDA 8  
 Db 101 FLOLEPDS 108

## RESULT 35

Q6PLU4 PRELIMINARY; PRT; 189 AA.  
 AC Q6PLU4;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 GN Name=COII;  
 OS Aegla sanlorenzo.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 OC Galatheoidea; Aegidae; Aegla.  
 OC NCBI\_TaxID=273846;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;  
 RT "Molecular systematics and biogeography of the southern South American  
 freshwater crabs Aegla (Decapoda: Anomura: Aegidae) using multiple  
 heuristic tree search approaches.";   
 RL Syst. Biol. 53:0-0(2004).  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 3 form the functional core of the enzyme complex. Subunit 2  
 transfers the electrons from cytochrome c via its binuclear copper  
 A center to the binuclear center of the catalytic subunit 1 (By  
 similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AY595780; AA08416.1; -  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR InterPro; IPR001505; Copper\_CUA.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR InterPro; IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PR0116; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CUA; 1.  
 DR Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1 1  
 FT NON\_TER 189 189  
 SQ SEQUENCE 189 AA; 21705 MW; 0DE6E2B6EF175110 CRC64;

Query Match 84.1%; Score 37; DB 2; Length 189;  
 Best Local Similarity 87.5%; Pred. No. 9.7;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEPDA 8

Db 101 FLOLEPDS 108

## RESULT 36

Q6PLU5 PRELIMINARY; PRT; 189 AA.  
 AC Q6PLU5;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 GN Name=COII;  
 OS Aegla sanlorenzo.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 OC Galatheoidea; Aegidae; Aegla.  
 OC NCBI\_TaxID=273846;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;  
 RT "Molecular systematics and biogeography of the southern South American  
 freshwater crabs Aegla (Decapoda: Anomura: Aegidae) using multiple  
 heuristic tree search approaches.";   
 RL Syst. Biol. 53:0-0(2004).  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 3 form the functional core of the enzyme complex. Subunit 2  
 transfers the electrons from cytochrome c via its binuclear copper  
 A center to the binuclear center of the catalytic subunit 1 (By  
 similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 c + 2 H(2)O.  
 CC -1- COFACTOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 DR EMBL; AY595779; AA08415.1; -  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR InterPro; IPR001505; Copper\_CUA.  
 DR InterPro; IPR008972; Cupredoxin.  
 DR InterPro; IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2\_TM; 1.  
 DR PRINTS; PR0116; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CUA; 1.  
 DR Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;  
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON\_TER 1 1  
 FT NON\_TER 189 189  
 SQ SEQUENCE 189 AA; 21747 MW; EB5CC5977BDP151B CRC64;

Query Match 84.1%; Score 37; DB 2; Length 189;  
 Best Local Similarity 87.5%; Pred. No. 9.7;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLOLEPDA 8  
 Db 101 FLOLEPDS 108

## RESULT 37

Q6PLU6 PRELIMINARY; PRT; 189 AA.  
 AC Q6PLU6;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 GN Name=COII;  
 OS Aegla jujuyana.  
 OG Mitochondrion.

CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 CC Galatheoidea; Aegliidae; Aegla.  
 OX NCBI\_TaxID=273843;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;  
 RT "Molecular systematics and biogeography of the southern South American  
 freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple  
 heuristic tree search approaches."  
 RL Syst. Biol. 53:0-0(2004).  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 3 form the functional core of the enzyme complex. Subunit 2  
 transfers the electrons from cytochrome c via its binuclear copper  
 A center to the bimetallic center of the catalytic subunit 1 (By  
 similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 c + 2 H(2)O.  
 CC -1- COPACITOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 CC EMBL; AY595777; AAT08414.1; -.  
 DR GO: 0005739; C:mitochondrion; IEA.  
 DR InterPro: IPR001505; Copper\_CuA.  
 DR InterPro: IPR008972; Cupredoxin.  
 DR InterPro: IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2; 1.  
 DR PRINTS; PR01166; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CuA; 1.  
 DR Kegg; E:electron transport; Inner membrane; Membrane; Mitochondrion;  
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON TER 1 189 1  
 FT NON TER 1 189 1  
 SQ SEQUENCE 189 AA; 21733 MW; C72D0CA6F1F7B791 CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 189;  
 Best Local Similarity 87.5%; Pred. No. 9.7;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 FLOLEPDA 8  
 Db 101 FLOLEPDS 108  
 RESULT 38  
 O6PLV0 PRELIMINARY; PRT; 189 AA.  
 AC O6PLV0; 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 RT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 GN Name=COII;  
 OS Aegla septentrionalis.  
 OC Mitochondrion.  
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 CC Galatheoidea; Aegliidae; Aegla.  
 OX NCBI\_TaxID=273844;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;  
 RT "Molecular systematics and biogeography of the southern South American  
 freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple  
 heuristic tree search approaches."  
 RL Syst. Biol. 53:0-0(2004).  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 3 form the functional core of the enzyme complex. Subunit 2

transfers the electrons from cytochrome c via its binuclear copper  
 A center to the bimetallic center of the catalytic subunit 1 (By  
 similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 c + 2 H(2)O.  
 CC -1- COPACITOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 CC EMBL; AY595776; AAT08412.1; -.  
 DR GO: 0005739; C:mitochondrion; IEA.  
 DR InterPro: IPR001505; Copper\_CuA.  
 DR InterPro: IPR008972; Cupredoxin.  
 DR InterPro: IPR002429; Cyt\_c\_ox\_2.  
 DR Pfam; PF00116; COX2; 1.  
 DR Pfam; PF02790; COX2; 1.  
 DR PRINTS; PR01166; CYCOXIDASEII.  
 DR ProDom; PD000131; Copper\_CuA; 1.  
 DR Kegg; E:electron transport; Inner membrane; Membrane; Mitochondrion;  
 KM Oxidoreductase; Respiratory chain; Transmembrane; Transport.  
 FT NON TER 1 189 1  
 FT NON TER 1 189 1  
 SQ SEQUENCE 189 AA; 21703 MW; E0CFA9E101F7A8B3 CRC64;  
 Query Match 84.1%; Score 37; DB 2; Length 189;  
 Best Local Similarity 87.5%; Pred. No. 9.7;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 FLOLEPDA 8  
 Db 101 FLOLEPDS 108  
 RESULT 39  
 O6PLV0 PRELIMINARY; PRT; 189 AA.  
 AC O6PLV0; 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 RT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Cytochrome oxidase subunit II (Fragment).  
 GN Name=COII;  
 OS Aegla scamosa.  
 OC Mitochondrion.  
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;  
 CC Galatheoidea; Aegliidae; Aegla.  
 OX NCBI\_TaxID=273843;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;  
 RT "Molecular systematics and biogeography of the southern South American  
 freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple  
 heuristic tree search approaches."  
 RL Syst. Biol. 53:0-0(2004).  
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory  
 chain that catalyzes the reduction of oxygen to water. Subunits 1-  
 3 form the functional core of the enzyme complex. Subunit 2  
 transfers the electrons from cytochrome c via its binuclear copper  
 A center to the bimetallic center of the catalytic subunit 1 (By  
 similarity).  
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 c + 2 H(2)O.  
 CC -1- COPACITOR: Copper A (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 inner membrane (By similarity).  
 CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.  
 CC EMBL; AY595774; AAT08410.1; -.  
 DR GO: 0005739; C:mitochondrion; IEA.  
 DR InterPro: IPR001505; Copper\_CuA.  
 DR InterPro: IPR008972; Cupredoxin.



```

DR InterPro: IPR002429; Cyt_c_ox_2.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR PRODOM; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 189 AA; 21701 MW; 8DBE60D0987A8772 CRC64;

Query Match      84.1%; Score 37; DB 2; Length 189;
Best Local Similarity 87.5%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLOLEFPA 8
Db 101 FLOLEFDS 108

RESULT 40
Q6PLV3
ID O6PLV3 PRELIMINARY; PRT; 189 AA.
AC O6PLV3;
DT 05-JUL-2004 (TRENBLREL. 27, Created)
DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN Name=COII;
OS Aegla obstepa.
OC Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Anomura;
OC Galatheoidea; Aegliidae; Aegla.
OX NCBI_TaxID=273842;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Losada M., Bond-Buckup G., Jara C.G., Crandall K.A.;
RT "Molecular systematics and biogeography of the southern South American
RT freshwater crabs Aegla (Decapoda: Anomura: Aegliidae) using multiple
RT heuristic tree search approaches.";
RL Syst. Biol. 53:0-0(2004).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
CC chain that catalyzes the reduction of oxygen to water. Subunits 1-
CC 3 form the functional core of the enzyme complex. Subunit 2
CC transfers the electrons from cytochrome c via its binuclear copper
CC A center to the binuclear center of the catalytic subunit 1 (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- COFACTOR: Copper A (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -1- SIMILARITY: Belongs to the cytochrome c oxidase subunit 2 family.
DR EMBL; AY595771; AAT08407.1; -.
DR GO; GO:0005739; C:mitochondrion; IBA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR02429; Cyt_c_ox_2.
DR Pfam; PF00116; COX2; 1.
DR Pfam; PF02790; COX2_TM; 1.
DR PRINTS; PR01166; CYCOXIDASEII.
DR PRODOM; PD000131; Copper_CuA; 1.
KW Copper; Electron transport; Inner membrane; Membrane; Mitochondrion;
KW Oxidoreductase; Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 189 AA; 21685 MW; F44E05F703289A73 CRC64;

Query Match      84.1%; Score 37; DB 2; Length 189;
Best Local Similarity 87.5%; Pred. No. 9.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 FLOLEFPA 8
Db 101 FLOLEFDS 108

Search completed: January 12, 2005, 20:14:06
Job time : 89 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:46:44 ; Search time 21.6 Seconds  
(without alignments)  
27.633 Million cell updates/sec

Title: US-09-870-216C-7  
Perfect score: 50  
Sequence: 1 FLWFEIDIV 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, AA:  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfilltest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	78.0	585	4 US-09-489-039A-13868	Sequence 13868, A
2	36	72.0	373	4 US-09-252-991A-24034	Sequence 24034, A
3	36	72.0	700	4 US-09-266-225D-6	Sequence 6, Appl1
4	35	70.0	543	4 US-09-540-236-2405	Sequence 2405, Ap
5	35	70.0	599	3 US-09-228-986-117	Sequence 117, App
6	35	70.0	599	4 US-10-101-464A-117	Sequence 117, App
7	35	70.0	816	4 US-10-101-464A-827	Sequence 827, App
8	35	70.0	951	4 US-09-328-352-4456	Sequence 4456, Ap
9	35	70.0	1240	4 US-10-101-464A-976	Sequence 976, App
10	34	68.0	262	4 US-09-252-991A-29018	Sequence 29018, A
11	33	66.0	86	4 US-09-270-767-60013	Sequence 60013, A
12	33	66.0	309	3 US-08-952-736A-10	Sequence 10, Appl
13	33	66.0	338	4 US-09-198-452A-1050	Sequence 1050, Ap
14	33	66.0	363	4 US-09-328-352-5561	Sequence 5561, Ap
15	33	66.0	516	2 US-08-676-166A-2	Sequence 2, Appl1
16	32	64.0	68	4 US-09-513-999C-4712	Sequence 4712, Ap
17	32	64.0	120	4 US-09-489-039A-8587	Sequence 8587, Ap
18	32	64.0	231	3 US-09-134-001C-4094	Sequence 4094, Ap
19	32	64.0	279	4 US-09-328-352-5581	Sequence 5581, Ap
20	32	64.0	308	4 US-09-328-352-7093	Sequence 7093, Ap
21	32	64.0	323	4 US-09-270-767-31938	Sequence 31938, A
22	32	64.0	328	4 US-09-583-110-4999	Sequence 4999, Ap
23	32	64.0	350	1 US-08-415-751-43	Sequence 43, Appl
24	32	64.0	419	4 US-09-489-039A-12769	Sequence 12769, A
25	32	64.0	480	4 US-09-107-532A-6160	Sequence 6160, Ap
26	32	64.0	481	3 US-08-617-785-8	Sequence 8, Appl1
27	32	64.0	481	4 US-09-817-464-8	Sequence 8, Appl1

28	32	64.0	511	1 US-08-220-151-17	Sequence 17, Appl
29	32	64.0	511	1 US-08-413-118-17	Sequence 17, Appl
30	32	64.0	511	3 US-08-473-446-17	Sequence 17, Appl
31	32	64.0	575	4 US-09-248-766A-17643	Sequence 17643, A
32	32	64.0	754	4 US-09-252-991A-33133	Sequence 33133, A
33	32	64.0	828	4 US-10-101-464A-934	Sequence 934, App
34	32	64.0	867	4 US-08-617-785-4	Sequence 4, Appl1
35	32	64.0	867	4 US-09-817-464-4	Sequence 4, Appl1
36	32	64.0	915	1 US-08-453-862-2	Sequence 2, Appl1
37	32	64.0	915	2 US-08-452-734A-2	Sequence 2, Appl1
38	32	64.0	915	3 US-08-617-785-12	Sequence 12, Appl
39	32	64.0	915	3 US-08-176-401B-2	Sequence 4, Appl1
40	32	64.0	915	4 US-09-817-464-12	Sequence 12, Appl
41	32	64.0	915	5 PCT-US94-14989-2	Sequence 2, Appl1
42	32	64.0	922	3 US-08-617-785-14	Sequence 14, Appl
43	32	64.0	922	4 US-09-817-464-14	Sequence 14, Appl
44	31	62.0	63	4 US-09-248-786A-21418	Sequence 21418, A
45	31	62.0	63	4 US-09-513-999C-5288	Sequence 5288, Ap

ALIGNMENTS

```
RESULT 1
US-09-489-039A-13868
; Sequence 13868, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13868
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13868

Query Match      78.0%; Score 39; DB 4; Length 585;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      3 WFEIDIV 9
Db      273 WFEIDIV 279

RESULT 2
US-09-252-991A-24034
; Sequence 24034, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24034
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24034
```

Query Match 72.0%; Score 36; DB 4; Length 373;  
Best Local Similarity 55.6%; Pred. No. 85;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
Db 307 YLWVELDLV 315

RESULT 3  
US-09-266-225D-6  
Sequence 6, Application US/09266225D  
Patent No. 6573364  
GENERAL INFORMATION:  
APPLICANT: Nandabalan, Krishan  
APPLICANT: Kingsmore, Stephen  
APPLICANT: Tchervnev, Velizar  
TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak  
TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-  
TITLE OF INVENTION: Interacting Proteins  
FILE REFERENCE: 15966-523  
CURRENT APPLICATION NUMBER: US/09/266,225D  
CURRENT FILING DATE: 1999-03-10  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 700  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-266-225D-6

Query Match 72.0%; Score 36; DB 4; Length 700;  
Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLWFEIDI 8  
Db 612 FLWFENDM 619

RESULT 4  
US-09-540-236-2405  
Sequence 2405, Application US/09540236  
Patent No. 6673910  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709,2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 2405  
LENGTH: 543  
TYPE: PRT  
ORGANISM: M. catarrhalis  
US-09-540-236-2405

Query Match 70.0%; Score 35; DB 4; Length 543;  
Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
Db 308 FLWFEIDLI 316

RESULT 5  
US-09-228-986-117  
Sequence 117, Application US/09228986  
Patent No. 6359198  
GENERAL INFORMATION:

APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, Niels  
TITLE OF INVENTION: Compositions Isolated from Plant Cells  
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
FILE REFERENCE: 11000/1020  
CURRENT APPLICATION NUMBER: US/09/228,986  
CURRENT FILING DATE: 1999-01-12  
NUMBER OF SEQ ID NOS: 130  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 117  
LENGTH: 599  
TYPE: PRT  
ORGANISM: Eucalyptus grandis  
US-09-228-986-117

Query Match 70.0%; Score 35; DB 3; Length 599;  
Best Local Similarity 83.3%; Pred. No. 2.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFEID 7  
Db 33 LWFEVD 38

RESULT 6  
US-10-101-464A-117  
Sequence 117, Application US/10101464A  
Patent No. 6768041  
GENERAL INFORMATION:  
APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, Nicolaas  
APPLICANT: Higgins, Colleen M.  
TITLE OF INVENTION: Compositions Isolated from Plant Cells  
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
FILE REFERENCE: 11000,1020c2  
CURRENT APPLICATION NUMBER: US/10/101,464A  
CURRENT FILING DATE: 2002-03-18  
PRIOR APPLICATION NUMBER: 09/704,302  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 09/228,986  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/162,866  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: PCT/US00/00724  
PRIOR FILING DATE: 2000-01-11  
NUMBER OF SEQ ID NOS: 989  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 117  
LENGTH: 599  
TYPE: PRT  
ORGANISM: Eucalyptus grandis  
US-10-101-464A-117

Query Match 70.0%; Score 35; DB 4; Length 599;  
Best Local Similarity 83.3%; Pred. No. 2.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFEID 7  
Db 33 LWFEVD 38

RESULT 7  
US-10-101-464A-827  
Sequence 827, Application US/10101464A  
Patent No. 6768041  
GENERAL INFORMATION:  
APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, Nicolaas  
APPLICANT: Higgins, Colleen M.  
TITLE OF INVENTION: Compositions Isolated from Plant Cells  
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
FILE REFERENCE: 11000,1020c2

```

; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 827
; LENGTH: 816
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-827

Query Match
Best Local Similarity 70.0%; Score 35; DB 4; Length 816;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMFID 7
Db 250 LMFVD 255

RESULT 8
US-09-328-352-4456
; Sequence 4456, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4456
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4456

Query Match
Best Local Similarity 70.0%; Score 35; DB 4; Length 951;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMFID 9
Db 801 VWFVDV 808

RESULT 9
US-10-101-464A-976
; Sequence 976, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
```

```

; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 976
; LENGTH: 1240
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-976

Query Match
Best Local Similarity 70.0%; Score 35; DB 4; Length 1240;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMFID 7
Db 674 LMFVD 679

RESULT 10
US-09-252-991A-29018
; Sequence 29018, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29018
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29018

Query Match
Best Local Similarity 66.0%; Score 34; DB 4; Length 262;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMFID 7
Db 110 LMFVD 115

RESULT 11
US-09-270-767-60013
; Sequence 60013, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 60013
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-60013

Query Match
Best Local Similarity 66.0%; Score 33; DB 4; Length 86;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 FLMFEID 7  
Db 78 FLMFDFD 84

RESULT 12  
US-08-952-736A-10  
; Sequence 10, Application US/08952736A  
; Patent No. 6320026  
; GENERAL INFORMATION:

APPLICANT: Cell Growth Inhibitory Factor  
TITLE OF INVENTION: 10  
NUMBER OF SEQUENCES: 10  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/952,736A  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 309 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-952-736A-10

Query Match 66.0%; Score 33; DB 3; Length 309;  
Best Local Similarity 71.4%; Pred. No. 2.4e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEID 7  
Db 29 FLMFDFD 35

RESULT 13  
US-09-198-452A-1050  
; Sequence 1050, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
APPLICANT: Griffiths, R.  
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
TITLE OF INVENTION: and treatment of infection  
FILE REFERENCE: 9710-003-999  
CURRENT APPLICATION NUMBER: US/09/198,452A  
CURRENT FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 6849  
SEQ ID NO 1050  
LENGTH: 338  
TYPE: PRT  
ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-1050

Query Match 66.0%; Score 33; DB 4; Length 338;  
Best Local Similarity 71.4%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEID 7  
Db 115 FLMFDFD 121

RESULT 14  
US-09-328-352-5961  
; Sequence 5961, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 5961  
LENGTH: 363  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-5961

Query Match 66.0%; Score 33; DB 4; Length 363;  
Best Local Similarity 71.4%; Pred. No. 2.8e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEID 7  
Db 133 FLMFDFD 139

RESULT 15  
US-08-676-166A-2  
; Sequence 2, Application US/08676166A  
; Patent No. 5955270  
; GENERAL INFORMATION:  
APPLICANT: Radford, Alan  
TITLE OF INVENTION: EXPLOITATION OF THE CELLULOSE COMPLEX OF  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
STREET: Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/676,166A  
FILING DATE: 15-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1321-1-002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 516 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-676-166A-2

Query Match 66.0%; Score 33; DB 2; Length 516;  
Best Local Similarity 57.1%; Pred. No. 4.1e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMFDFD 8  
Db 139 LMFDFD 145

RESULT 16

US-09-513-999C-4712  
; Sequence 4712, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duciart, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59, US2, REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 4712  
; LENGTH: 68  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -22...-1  
; OTHER INFORMATION: score 5.9  
; OTHER INFORMATION: seq L1LWFMHDCVSS/IL  
US-09-513-999C-4712

Query Match 64.0%; Score 32; DB 4; Length 68;

Best Local Similarity 62.5%; Pred. No. 73; Indels 0; Gaps 0;

Qy 2 LWFEDIV 9  
Db 13 LWFHDCV 20

RESULT 17  
US-09-489-8587  
; Sequence 8587, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709, 2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 8587  
; LENGTH: 120  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-8587

Query Match 64.0%; Score 32; DB 4; Length 120;

Best Local Similarity 37.5%; Pred. No. 1.3e+02; Indels 0; Gaps 0;

Qy 2 LWFEDIV 9  
Db 34 MWYEDVI 41

RESULT 18  
US-09-134-001C-4094  
; Sequence 4094, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4094  
; LENGTH: 231  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4094

Query Match 64.0%; Score 32; DB 3; Length 231;

Best Local Similarity 57.1%; Pred. No. 2.7e+02; Indels 0; Gaps 0;

Qy 3 WFEIDIV 9  
Db 108 WFDIVV 114

RESULT 19  
US-09-328-352-5581  
; Sequence 5581, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 5581  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-5581

Query Match 64.0%; Score 32; DB 4; Length 279;

Best Local Similarity 55.6%; Pred. No. 3.2e+02; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
Db 34 FWFAGDIV 42

RESULT 20  
US-09-328-352-7093  
; Sequence 7093, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7093  
; LENGTH: 308  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7093

Query Match 64.0%; Score 32; DB 4; Length 308;

Best Local Similarity 62.5%; Pred. No. 3.6e+02; Indels 0; Gaps 0;

Qy 1 FLWFEIDI 8

Db :|||: ||  
67 YLMFKPDI 74

RESULT 21  
US-09-270-767-31938  
; Sequence 31938, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 31938  
; LENGTH: 323  
; TYPE: PRT  
; ORGANISM: *Drosophila melanogaster*  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-31938

Query Match 64.0%; Score 32; DB 4; Length 323;  
Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLWFE 5  
|||:  
Db 223 FLWFE 227

RESULT 22  
US-09-583-110-4999  
; Sequence 4999, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to *Streptococcus*  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 4999  
; LENGTH: 328  
; TYPE: PRT  
; ORGANISM: *Streptococcus pneumoniae*  
US-09-583-110-4999

Query Match 64.0%; Score 32; DB 4; Length 328;  
Best Local Similarity 83.3%; Pred. No. 3.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LWFED 7  
|||:  
Db 73 LWFED 78

RESULT 23  
US-08-415-751-43  
; Sequence 43, Application US/08415751  
; Patent No. 5643772  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES

APPLICANT: NELSON, RICHARD, C.  
; TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-  
; TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA  
; TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID  
; TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND  
; TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: PHILLIPS, MOORE, LEMPLO & FINLEY  
; STREET: 365 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage  
; COMPUTER: PC  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/415,751  
; FILING DATE: 03-Apr-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/071,880  
; FILING DATE: June 1, 1993  
; APPLICATION NUMBER: 07/891,301  
; FILING DATE: May 29, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hana Dolezalova  
; REGISTRATION NUMBER: 30,518  
; REFERENCE/DOCKET NUMBER: 480.19-2 (HHD)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-1677  
; TELEFAX: (415) 324-1678  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 350 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: *Cryptosporidium parvum*  
; FEATURE:  
; NAME/KEY: Positions coded by nonsense codons are  
; NAME/KEY: Identified as Xaa.  
US-08-415-751-43

Query Match 64.0%; Score 32; DB 1; Length 350;  
Best Local Similarity 66.7%; Pred. No. 4.1e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
|||:  
Db 236 FLWFEIDIV 244

RESULT 24  
US-09-489-039A-12789  
; Sequence 12789, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *KLEBSIELLA*  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29



NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO: 12789  
LENGTH: 419  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12789

Query Match 64.0%; Score 32; DB 4; Length 419;  
Best Local Similarity 50.0%; Pred. No. 5e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEIDI 8  
Db 45 FIMFEDEV 52

RESULT 25  
US-09-107-532A-6160  
Sequence 6160, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESSES:  
ADDRESS: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariellello, Pamela Denise  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 6160:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 480 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...480  
SEQUENCE DESCRIPTION: SEQ ID NO: 6160:  
US-09-107-532A-6160

Query Match 64.0%; Score 32; DB 4; Length 480;  
Best Local Similarity 62.5%; Pred. No. 5.8e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 FLMFEIDI 8  
Db 1 FLMFEIDI 8

Db 99 FLMFEIDI 106

RESULT 26  
US-08-617-785-8  
Sequence 8, Application US/08617785E  
Patent No. 6228610  
GENERAL INFORMATION:  
APPLICANT: Flor, Peter J.  
APPLICANT: Kuhn, Rainer  
APPLICANT: Lindaur, Kristen  
APPLICANT: Puttner, Irene  
APPLICANT: Knopfel, Thomas  
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,  
FILE REFERENCE: 4-19679/A/PCT  
CURRENT APPLICATION NUMBER: US/08/617,785E  
CURRENT FILING DATE: 1996-03-19  
EARLIER APPLICATION NUMBER: PCT/EP94/02991  
EARLIER FILING DATE: 1994-09-07  
EARLIER APPLICATION NUMBER: EPO 9416553.7  
EARLIER FILING DATE: 1994-08-19  
EARLIER APPLICATION NUMBER: EPO 93810663.0  
EARLIER FILING DATE: 1993-09-20  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 8  
LENGTH: 481  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-617-785-8

Query Match 64.0%; Score 32; DB 3; Length 481;  
Best Local Similarity 57.1%; Pred. No. 5.8e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEIDI 7  
Db 293 FIMFEVD 299

RESULT 27  
US-09-817-464-8  
Sequence 8, Application US/09817464  
Patent No. 6515107  
GENERAL INFORMATION:  
APPLICANT: Flor, Peter J.  
APPLICANT: Kuhn, Rainer  
APPLICANT: Puttner, Irene  
APPLICANT: Knopfel, Thomas  
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,  
FILE REFERENCE: 4-19679/A/PCT  
CURRENT APPLICATION NUMBER: US/09/817,464  
CURRENT FILING DATE: 2001-03-26  
EARLIER APPLICATION NUMBER: US/08/617,785  
EARLIER FILING DATE: 1996-03-19  
EARLIER APPLICATION NUMBER: EPO 9416553.7  
EARLIER FILING DATE: 1994-08-19  
EARLIER APPLICATION NUMBER: EPO 93810663.0  
EARLIER FILING DATE: 1993-09-20  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 8  
LENGTH: 481  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-817-464-8

Query Match 64.0%; Score 32; DB 4; Length 481;  
Best Local Similarity 57.1%; Pred. No. 5.8e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEID 7  
Db 293 FLMFEVD 299

## RESULT 28

US-08-220-151-17  
Sequence 17, Application US/08220151

Patent No. 5529780  
GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
APPLICANT: Limbach, Keith J.  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC AND GD AND USES THEREFOR  
NUMBER OF SEQUENCES: 91  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/220,151  
FILING DATE: 30-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066 CURTMS  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 511 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-220-151-17

Query Match 64.0%; Score 32; DB 1; Length 511;  
Best Local Similarity 71.4%; Pred. No. 6.1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEID 7  
Db 299 FLMFEVD 305

## RESULT 29

US-08-413-118-17  
Sequence 17, Application US/08413118

Patent No. 5688920  
GENERAL INFORMATION:  
APPLICANT: PAOLETTI, ENZO  
APPLICANT: LIMBACH, KEITH J.  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC, AND GD AND USES THEREFOR  
NUMBER OF SEQUENCES: 128  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.  
STREET: 530 FIFTH AVENUE, 25TH FLOOR  
CITY: NEW YORK

STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/413,118  
FILING DATE: 23-MAR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/220,151  
FILING DATE: 30-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: FROMMER, WILLIAM S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2670  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 511 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-413-118-17

Query Match 64.0%; Score 32; DB 1; Length 511;  
Best Local Similarity 71.4%; Pred. No. 6.1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEID 7  
Db 299 FLMFEVD 305

RESULT 30  
US-08-473-446-17  
Sequence 17, Application US/08473446

Patent No. 6017542  
GENERAL INFORMATION:  
APPLICANT: PAOLETTI, ENZO  
APPLICANT: LIMBACH, KEITH J.  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC, AND GD AND USES THEREFOR  
NUMBER OF SEQUENCES: 128  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.  
STREET: 530 FIFTH AVENUE, 25TH FLOOR  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,446  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/413,118  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: FROMMER, WILLIAM S.  
REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454310-2670  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 511 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-473-446-17

Query Match 64.0%; Score 32; DB 3; Length 511;  
Best Local Similarity 71.4%; Pred. No. 6.1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWFEID 7  
Db 299 FWMFEDD 305

RESULT 31  
US-09-248-796A-17643  
Sequence 17643, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstein et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.1132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
PRIOR FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 17643  
LENGTH: 575  
TYPE: PRT  
ORGANISM: Candida albicans  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (53) (57) (59)  
OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown  
US-09-248-796A-17643

Query Match 64.0%; Score 32; DB 4; Length 575;  
Best Local Similarity 42.9%; Pred. No. 7e+02;  
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLWFEID 7  
Db 416 YLMYQVD 422

RESULT 32  
US-09-252-991A-33133  
Sequence 33133, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.1136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 33133  
LENGTH: 754  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-33133

Query Match 64.0%; Score 32; DB 4; Length 754;  
Best Local Similarity 100.0%; Pred. No. 9.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WFEID 7  
Db 235 WFEID 239

RESULT 33  
US-10-101-464A-934  
Sequence 934, Application US/10101464A  
Patent No. 6768041  
GENERAL INFORMATION:  
APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, Nicolaas  
TITLE OF INVENTION: Compositions Isolated from Plant Cells  
FILE REFERENCE: 11000.1020c2  
CURRENT APPLICATION NUMBER: US/10/101,464A  
PRIOR FILING DATE: 2002-03-18  
PRIOR APPLICATION NUMBER: 09/704,302  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 09/228,986  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/162,866  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: PCT/US00/00724  
PRIOR FILING DATE: 2000-01-11  
NUMBER OF SEQ ID NOS: 989  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 934  
LENGTH: 828  
TYPE: PRT  
ORGANISM: Pinus radiata  
US-10-101-464A-934

Query Match 64.0%; Score 32; DB 4; Length 828;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLWFE 5  
Db 292 FLWFE 296

RESULT 34  
US-08-617-785-4  
Sequence 4, Application US/08617785E  
Patent No. 6228610  
GENERAL INFORMATION:  
APPLICANT: Flor, Peter J.  
APPLICANT: Kuhn, Rantier  
APPLICANT: Lindaur, Kristen  
APPLICANT: Putner, Irene  
APPLICANT: Knopfel, Thomas  
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,  
FILE REFERENCE: 4-19679/A/PCT  
CURRENT APPLICATION NUMBER: US/08/617,785E  
PRIOR FILING DATE: 1996-03-19  
PRIOR APPLICATION NUMBER: PCT/EP94/02991  
PRIOR FILING DATE: 1994-09-07  
PRIOR APPLICATION NUMBER: EPO 9416553.7  
PRIOR FILING DATE: 1994-08-19

EARLIER APPLICATION NUMBER: EPO 93810663.0  
EARLIER FILING DATE: 1993-09-20  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 867  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-617-785-4

Query Match 64.0%; Score 32; DB 3; Length 867;  
Best Local Similarity 57.1%; Pred. No. 1.1e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEID 7  
DB 663 FIVEGVD 669

RESULT 35  
US-09-817-464-4  
Sequence 4, Application US/09817464  
GENERAL INFORMATION:  
APPLICANT: Pior, Peter J.  
APPLICANT: Kuhn, Rainer  
APPLICANT: Lindner, Irene  
APPLICANT: Knopfel, Thomas  
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,  
FILE REFERENCE: 4-19679/A/PCT  
CURRENT FILING DATE: 2001-03-26  
EARLIER APPLICATION NUMBER: US/09/817,464  
EARLIER FILING DATE: 1996-03-19  
EARLIER APPLICATION NUMBER: EPO 9416553.7  
EARLIER FILING DATE: 1994-08-19  
EARLIER APPLICATION NUMBER: EPO 93810663.0  
EARLIER FILING DATE: 1993-09-20  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 867  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-817-464-4

Query Match 64.0%; Score 32; DB 4; Length 867;  
Best Local Similarity 57.1%; Pred. No. 1.1e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEID 7  
DB 663 FIVEGVD 669

RESULT 36  
US-08-453-862-2  
Sequence 2, Application US/08453862  
GENERAL INFORMATION:  
APPLICANT: Segerson, Thomas P.  
APPLICANT: Kinzie, J. Mark  
APPLICANT: Mulvihill, Eileen R.  
APPLICANT: Saugstad, Julie A.  
TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco

STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/453,862  
FILING DATE: 30-MAY-1995  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,401  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 13952-18-2  
TELEPHONE: 206-467-9600  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 915 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-453-862-2

Query Match 64.0%; Score 32; DB 1; Length 915;  
Best Local Similarity 57.1%; Pred. No. 1.1e+03;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEID 7  
DB 718 FIVEGVD 724

RESULT 37  
US-08-452-734A-2  
Sequence 2, Application US/08452734A  
Patent No. 5831047  
GENERAL INFORMATION:  
APPLICANT: Segerson, Thomas P.  
APPLICANT: Kinzie, J. Mark  
APPLICANT: Mulvihill, Eileen R.  
APPLICANT: Saugstad, Julie A.  
APPLICANT: Westbrock, Gary L.  
TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,734A  
FILING DATE: 30-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,401  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990

```
REFERENCE/DOCKET NUMBER: 13952-18-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 915 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-452-734A-2

Query Match
Best Local Similarity 64.0%; Score 32; DB 2; Length 915;
Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEID 7
Db 718 FLMFEVD 724

RESULT 38
US-08-617-785-12
Sequence 12, Application US/08617785E
Patent No. 6228610
GENERAL INFORMATION:
APPLICANT: Flor, Peter J.
APPLICANT: Kuhn, Rainer
APPLICANT: Lindaur, Kristen
APPLICANT: Putner, Irene
APPLICANT: Knopfel, Thomas
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
FILE REFERENCE: 4-19679/A/PCT
CURRENT APPLICATION NUMBER: US/08/617,785E
EARLIER FILING DATE: 1996-03-19
EARLIER APPLICATION NUMBER: PCT/EP94/02991
EARLIER FILING DATE: 1994-09-07
EARLIER APPLICATION NUMBER: EPO 9416553.7
EARLIER FILING DATE: 1994-08-19
EARLIER APPLICATION NUMBER: EPO 93810663.0
EARLIER FILING DATE: 1993-09-20
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 12
LENGTH: 915
TYPE: PRT
ORGANISM: Homo sapiens
US-08-617-785-12

Query Match
Best Local Similarity 64.0%; Score 32; DB 3; Length 915;
Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEID 7
Db 718 FLMFEVD 724

RESULT 39
US-08-176-401B-2
Sequence 2, Application US/08176401B
Patent No. 6274330
GENERAL INFORMATION:
APPLICANT: Segerson, Thomas P.
APPLICANT: Kinzie, J. Mark
APPLICANT: Mulvihill, Bileen R.
APPLICANT: Saugstad, Julie A.
APPLICANT: Westbrook, Gary L.
TITLE OF INVENTION: L-AP4 SENSITIVE GLUTAMATE RECEPTORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
```

```
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,401B
FILING DATE: 30-DECEMBER-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 13952-18-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 915 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-401B-2

Query Match
Best Local Similarity 64.0%; Score 32; DB 3; Length 915;
Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEID 7
Db 718 FLMFEVD 724

RESULT 40
US-09-817-464-12
Sequence 12, Application US/09817464
Patent No. 6515107
GENERAL INFORMATION:
APPLICANT: Flor, Peter J.
APPLICANT: Kuhn, Rainer
APPLICANT: Lindaur, Kristen
APPLICANT: Putner, Irene
APPLICANT: Knopfel, Thomas
TITLE OF INVENTION: Human Metabotropic Glutamate Receptor Subtypes (HMR4,
FILE REFERENCE: 4-19679/A/PCT
CURRENT APPLICATION NUMBER: US/09/817,464
EARLIER FILING DATE: 2001-03-26
EARLIER APPLICATION NUMBER: US/08/617,785
EARLIER FILING DATE: 1996-03-19
EARLIER APPLICATION NUMBER: EPO 9416553.7
EARLIER FILING DATE: 1994-08-19
EARLIER APPLICATION NUMBER: EPO 93810663.0
EARLIER FILING DATE: 1993-09-20
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 12
LENGTH: 915
TYPE: PRT
ORGANISM: Homo sapiens
US-09-817-464-12

Query Match
Best Local Similarity 64.0%; Score 32; DB 4; Length 915;
Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEID 7
```

Thu Jan 13 08:15:10 2005

us-09-870-216c-7.ra1

Db 718 FIMGVD 724

Search completed: January 12, 2005, 20:17:40  
Job time : 22.6 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:37:43 ; Search time 201.4 Seconds

(without alignments)  
16.031 Million cell updates/sec

Title: US-09-870-216C-7

Perfect score: 50

Sequence: 1 FLWFEIDIV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A\_Geneseq\_23Sep04.\*

1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	9	5	ABO8363 Synthetic
2	50	100.0	9	7	ABR8215 Human ant
3	50	100.0	352	5	ABO8368 Human can
4	39	78.0	585	7	ABO67351 Klebsiell
5	39	78.0	587	6	AAO30104 Yersinia
6	39	78.0	1717	4	ABR20498 Oestrin
7	38	76.0	589	6	ABM69298 Photorhab
8	37	74.0	150	5	ABR49338 Listeria
9	37	74.0	273	5	ABP64842 Human pro
10	37	74.0	528	5	ABP64842 Human pro
11	37	74.0	587	6	ABM70700 Photorhab
12	36	72.0	269	5	ABR53431 Lactococc
13	36	72.0	280	7	ABO75288 Pseudomon
14	36	72.0	373	7	ABO75288 Pseudomon
15	36	72.0	410	4	ABG01159 Novel hum
16	36	72.0	591	6	ABM68497 Photorhab
17	36	72.0	652	6	ABO14689 Novel hum
18	36	72.0	700	6	ABO14690 Novel hum
19	36	72.0	700	7	ADJ14125 Human src
20	36	72.0	700	7	ADJ14125 Human src
21	36	72.0	1097	4	ABG01151 Human hea
22	35	70.0	217	5	ADJ17021 Human NOV
23	35	70.0	232	7	ADC86399 Human GPC
24	35	70.0	307	4	AAJ71459 Human Olf
25	35	70.0	315	4	AAJ72215 Human Olf

26	35	70.0	315	5	ABG79345 Human GPC
27	35	70.0	315	5	ADJ16656 Human NOV
28	35	70.0	315	6	ABR01643 Human G P
29	35	70.0	315	8	ADN42310 Human nov
30	35	70.0	326	5	ABP95915 Human GPC
31	35	70.0	326	5	AAU95516 Human Olf
32	35	70.0	326	7	ADC85959 Human GPC
33	35	70.0	326	7	ADC85993 Human GPC
34	35	70.0	326	8	ADG83292 Human Olf
35	35	70.0	379	8	ADM42835 Rat odour
36	35	70.0	543	8	ADL04719 M. catarr
37	35	70.0	599	8	ABR25149 Eucalyptu
38	35	70.0	816	3	ABR25508 Eucalyptu
39	35	70.0	951	6	ADA33169 Acinetoba
40	34	68.0	37	4	AAH88375 Human imm
41	34	68.0	73	4	AAH89032 Human imm
42	34	68.0	195	3	AAJ21340 Arabidops
43	34	68.0	209	3	AAI18245 Plasmodi
44	34	68.0	215	3	AAJ38094 Arabidops
45	34	68.0	215	3	AAJ21339 Arabidops

## ALIGNMENTS

RESULT 1	
ABBO8363	
ID	ABBO8363 standard; protein; 9 AA.
AC	
XX	ABBO8363;
XX	
DT	07-MAY-2002 (first entry)
XX	
DE	Synthetic epitope 3 of human cancer antigen eIF3.
XX	
XX	Human; melanoma antigen eukaryotic initiation factor 3; eIF3;
KW	ovarian cancer; MHC; cytosolic; immunomodulator; immune effector cell
KW	anti-cancer; vaccine.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	Domain
FT	1 /note= "HLA-2 binding residue"
FT	2 /note= "HLA-2 binding residue"
FT	3..8 /note= "T-cell receptor (TCR) binding domain"
FT	9 /note= "HLA-2 binding residue"
FT	
FT	
XX	
XX	WO200192307-A2.
XX	
PD	06-DEC-2001.
XX	
PP	30-MAY-2001; 2001WO-US017456.
XX	
PR	31-MAY-2000; 2000US-0209391P.
PR	17-AUG-2000; 2000US-0226258P.
PR	20-DEC-2000; 2000US-0257008P.
XX	
PA	(GENZ ) GENZYME CORP.
XX	
PI	Nicolette CA;
XX	
DR	WPI; 2002-139606/18.
DR	N-PSDB; ABA97214.
XX	
PT	New therapeutic compounds useful against human ovarian cancer, for
PT	modulating immune response in a subject, and for generating antibodies
PT	that specifically recognize and bind to these molecules.
XX	
PS	Claim 29; Page 59; 68pp; English





PT modulating immune response in a subject, and for generating antibodies  
PT that specifically recognize and bind to these molecules.  
PS Claim 8; Page; 68pp; English.  
XX  
CC The invention relates to novel therapeutic compounds, that are designed  
CC to enhance binding to MHC molecules and to enhance immunoregulatory  
CC properties relative to their natural counterparts. The activity of the  
CC compounds of the invention may be described as cytostatic and  
CC immunomodulatory. The compounds are useful against human ovarian cancer,  
CC for modulating immune response in a subject, and for generating  
CC antibodies that specifically recognize and bind to these molecules.  
CC Compositions comprising the compounds are useful as components of anti-  
CC cancer vaccines and to expand immune effector cells that are specific for  
CC cells characterised by expression of antigen ERF3 (melanoma antigen  
CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
CC a detectable agent may be used in diagnostic procedures, such as in the  
CC detection and purification of antibodies, and as immunogens for  
CC production of antibodies. The polynucleotides can be used as primers for  
CC detecting genes or gene transcripts expressed in APC to confirm  
CC transduction of the polynucleotides into host cells. The current sequence  
CC represents the human cancer antigen elf3 variant 3 amino acid sequence.  
CC Note: This sequence is not present in the specification, but may be  
CC created from the sequence of the wild-type human cancer antigen elf3  
CC sequence given in ABB08360  
CC  
SQ Sequence 352 AA;  
XX  
Query Match 100.0%; Score 50; DB 5; Length 352;  
Best Local Similarity 100.0%; Pred. No. 0.94;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLWFEIDIV 9  
Db 242 FLWFEIDIV 250  
XX  
RESULT 4  
AB067351  
ID AB067351 standard; protein; 585 AA.  
XX  
AC AB067351;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Klebsiella pneumoniae polypeptide seqid 13868.  
XX  
XX  
KW Recombinant expression vector; transcription regulatory element;  
KM Klebsiella pneumoniae protein; antibacterial; Vaccine.  
XX  
OS Klebsiella pneumoniae.  
XX  
PN US6610836-B1.  
XX  
PD 26-AUG-2003.  
XX  
PF 27-JAN-2000; 2000US-00489039.  
XX  
PR 29-JAN-1999; 99US-0117747P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Breton GL, Osborne M;  
XX  
DR WPI; 2003-895346/82.  
DR N-PSDB; ABD00922.  
XX  
PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for  
PT preparing a vaccine composition against Klebsiella pneumoniae.  
XX  
PS Disclosure; SEQ ID NO 13868; 932pp; English.  
XX  
CC The invention describes a new isolated nucleic acid encoding a Klebsiella

CC pneumoniae polypeptide. Also described are: a recombinant expression  
CC vector comprising the nucleic acid, operably linked to a transcription  
CC regulatory element; and a cell comprising the recombinant expression  
CC vector. The nucleic acid is useful for preparing a vaccine composition  
CC against Klebsiella pneumoniae. This is the amino acid sequence of a  
CC Klebsiella pneumoniae polypeptide of the invention  
XX  
SQ Sequence 585 AA;  
XX  
Query Match 78.0%; Score 39; DB 7; Length 585;  
Best Local Similarity 85.7%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 WFEIDIV 9  
Db 273 WFEIDIV 279  
XX  
RESULT 5  
AA030104  
ID AA030104 standard; protein; 587 AA.  
XX  
AC AA030104;  
XX  
DT 03-SEP-2003 (first entry)  
XX  
DE Yersinia pestis 2G5 mutant protein.  
XX  
KM Infection; tuberculosis; veterinary treatment; antimicrobial drug;  
KM vaccine; mutant; mulein.  
XX  
OS Yersinia pestis.  
OS Synthetic.  
XX  
PN WO2003044047-A2.  
XX  
PD 30-MAY-2003.  
XX  
PF 18-NOV-2002; 2002WO-GB005212.  
XX  
PR 19-NOV-2001; 2001GB-00027657.  
XX  
XX (UNIL ) LONDON SCHOOL HYGIENE & TROPICAL MEDICIN.  
XX  
PI Karlyshev AV, Wren BW;  
XX  
DR WPI; 2003-457586/43.  
DR N-PSDB; AAL60720.  
XX  
PT New virulence peptides encoded by a gene of Yersinia pseudotuberculosis,  
PT useful for treating or preventing a condition associated with infection  
PT by Yersinia, e.g. tuberculosis, particularly in veterinary treatment.  
XX  
XX Claim 4; Page 96-99; 130pp; English.  
XX  
XX The present invention relates to virulence genes of Yersinia species and  
XX proteins encoded by such genes. Sequences of the invention are useful in  
XX the manufacture of a medicament for treating or preventing a condition  
XX associated with infection by Yersinia e.g. tuberculosis particularly in  
XX veterinary treatment. They are useful in screening assays for identifying  
XX antimicrobial drugs. They are also used to prepare vaccines. The present  
XX sequence is Yersinia pestis 2G5 mutant protein  
XX  
SQ Sequence 587 AA;  
XX  
Query Match 78.0%; Score 39; DB 6; Length 587;  
Best Local Similarity 85.7%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 WFEIDIV 9  
Db 272 WFEIDIV 278

```

RESULT 6
AAB20498
ID AAB20498 standard; protein; 1717 AA.
XX
XX AAB20498;
AC
XX 06-AUG-2003 (revised)
DT 09-JUL-2001 (first entry)
DE Ostrinia nubilalis Bt toxin receptor.
XX
XX European corn borer; ECB; Bacillus thuringiensis; Bt toxin; receptor;
KM crystal protein; CryIA; biological control; insecticide; crop protection.
XX
OS Ostrinia nubilalis.
XX
XX Key Location/Qualifiers
FH Binding-site 1293..1462
FT /label=CryIA_binding_site
FT 1571..1589
FT Domain /label=Transmembrane_domain
XX
XX WO200136639-A2.
XX
XX 25-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-US031674.
XX
XX 18-NOV-1999; 99US-0166285P.
XX 21-SEP-2000; 2000US-0234099P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Flanagan RD, Machis JP, Meyer TE;
XX WPI; 2001-329223/34.
XX N-PSDB; AAF30933.
XX
XX New nucleic acid encoding Bt toxin receptor, useful for screening and
PT identifying new Bt toxin receptor ligands useful as new insecticidal
PT toxins.
XX
XX Claim 4(a); Page 57-61; 85pp; English.
XX
XX The present sequence of is that of the Bacillus thuringiensis (Bt) toxin
CC receptor of the European corn borer (ECB), Ostrinia nubilalis. The
CC sequence was deduced from cDNA isolated from an ECB larva midgut cDNA.
CC The receptor protein binds CryIA toxin, particularly CryIA(b) toxins.
CC The invention provides compositions and methods for modulating the
CC susceptibility of a cell to Bt toxins. The compositions include Bt toxin
CC receptor proteins from the lepidopteran insects ECB, corn earworm
CC (Heliothis zea) and fall armyworm (Spodoptera frugiperda). Nucleic acids
CC encoding the polypeptides, antibodies specific to the polypeptides, and
CC nucleic acid constructs for expressing the polypeptides in cells are also
CC provided. The methods of the invention are useful for investigating the
CC structure-function relationships of Bt toxin receptors, investigating
CC toxin-receptor interactions, elucidating the mode of action of Bt toxins,
CC screening and identifying novel Bt toxin receptor ligands including novel
CC insecticidal toxins, and designing and developing novel Bt toxin receptor
CC ligands. The methods are useful for managing Bt toxin resistance in plant
CC pests, and protecting plants against damage by plant pests. (Updated on
CC 06-AUG-2003 to correct OS field.)
XX
XX Sequence 1717 AA;
SQ

```

```

Query Match 78.0%; Score 39; DB 4; Length 1717;
Best Local Similarity 85.7%; Pred. No. 4.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 FLMFEID 7
DB 650 YLMFEID 656

```

```

RESULT 7
ABM69298
ID ABM69298 standard; protein; 589 AA.
XX
XX ABM69298;
AC
XX 20-NOV-2003 (first entry)
DT
DE Photorhabdus luminescens protein sequence #2395.
XX
XX Anticardial; fungicide; insecticide; polymorphism; genetic analysis;
KM detection; food; gene expression; plant; animal; microorganism; toxin;
KM antibiotic; biopesticide; virulence factor; disease model; plague;
KM whooping cough.
XX
XX Photorhabdus luminescens.
XX
XX WO200294867-A2.
XX
XX 28-NOV-2002.
XX
XX 07-FEB-2002; 2002WO-IB003040.
XX
XX 07-FEB-2001; 2001FR-00001659.
XX
XX (INSP ) INST PASTEUR.
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Duchaud E, Taourit S, Glaeser P, Frangeul L, Kunst F, Danchin A,
XX Buchrieser C;
XX WPI; 2003-148459/14.
XX
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
PT
XX Claim 2; SEQ ID NO 2395; 1205pp; French.
XX
XX The invention relates to the isolation of genes and their encoded
CC proteins from Photorhabdus luminescens. The isolated sequences are
CC sources of probes and primers for detecting the genome of P. luminescens
CC and related species; to study polymorphisms; for gene analysis and for
CC detection/amplification of the genes. Antibodies (Ab) raised against the
CC polypeptides encoded by the genes are used for detection/identification
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC carry a gene-containing vector are used to select compounds that
CC modulate, regulate, induce or inhibit expression of the genes in plants,
CC animals or microorganisms other than P. luminescens and are able to alter
CC response or sensitivity to toxins and antibiotics produced by P.
CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful.
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens proteins
XX
XX Sequence 589 AA;
SQ

```

```

Query Match 76.0%; Score 38; DB 6; Length 589;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 3 WFEIDIV 9
DB 273 WFEIDIV 279

```

RESULT 8  
ID ABB49338 standard; protein; 150 AA.  
XX  
AC ABB49338;  
XX  
DT 05-FEB-2002 (first entry)  
XX  
DE Listeria monocytogenes protein #2042.  
XX  
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
KW vitamin B12; bacterial infection; disease.  
XX  
OS Listeria monocytogenes.  
XX  
PN MO200177335-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 11-APR-2001; 2001WO-FR001118.  
XX  
PR 11-APR-2000; 2000FR-00004629.  
XX  
PA (INSP ) INST PASTEUR.  
XX  
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;  
PI Duesurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;  
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng B, Vazquez-Boland JA;  
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
PI Chakraborty T, Domann E, Hain T, Berche P, Chabrit A, Durant L;  
PI Perez-Diaz J, Bagueri F, Garcia Del Portillo F, Gomez-Lopez N;  
PI Meduendo E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J;  
PI Rose M, Voss H;  
XX  
DR WPI; 2002-010914/01.  
XX  
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment  
PT and prevention of Listeria and related bacterial infections, and related  
PT polypeptides.  
XX  
PS Claim 6; SEQ ID NO 2043; 192bp; French.  
XX  
CC The present invention relates to the genome sequence of Listeria  
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of  
CC it are useful for selecting probes and primers for detecting genes in L.  
CC monocytogenes and related organisms, and for studying genetic  
CC polymorphisms and other genomes. The present sequence is a protein  
CC encoded by the genome sequence of the present invention. Proteins  
CC expressed from the genome sequence are useful for raising specific  
CC antibodies, identification of L. monocytogenes and related organisms, and  
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin  
CC B12. The genome sequence and proteins encoded by it are also useful for  
CC selecting compounds that regulate gene expression and cell replication  
CC and modulate L. monocytogenes-related diseases. In addition, the genome  
CC sequence and proteins encoded by it are useful in pharmaceutical and  
CC vaccines compositions for the treatment or prevention of infections by L.  
CC monocytogenes and related organisms. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 150 AA;  
XX  
Query Match 74.0%; Score 37; DB 5; Length 150;  
Best Local Similarity 77.8%; Pred. No. 80;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
1 FLWFEIDIV 9  
| | | | | | | | | |  
116 FKWFEIDKV 124

ABP64842  
ID ABB64842 standard; protein; 273 AA.  
XX  
AC ABB64842;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human protein SEQ ID 502.  
XX  
KW Human; expressed sequence tag; EST; haematopoietic disorder;  
KW central nervous system disease; viral infection;  
KW peripheral nervous system disease; non-healing wound; infectious disease;  
KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;  
KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;  
KW antileptic; antiinflammatory; immunosuppressive; neuroprotective;  
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;  
KW immunostimulant; cerebroprotective.  
XX  
OS Homo sapiens.  
XX  
PN MO200259260-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 16-NOV-2001; 2001WO-US042950.  
XX  
PR 17-NOV-2000; 2000US-00714936.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX  
DR WPI; 2002-590824/63.  
DR  
DR N-PSDB; ABQ99428.  
XX  
PT New isolated polynucleotide, useful in research, diagnostic or  
PT therapeutic methods, e.g. preventing or treating disorders involving  
PT aberrant protein expression or biological activity.  
XX  
PS Claim 20; SEQ ID NO 502; 394bp; English.  
XX  
CC The present invention relates to novel human coding sequences (ABQ99268-  
CC ABQ99608) and proteins (ABP6482-ABP65022). The sequences are useful in  
CC therapeutic, diagnostic and research methods. The polynucleotides may be  
CC used in the field of molecular biology as hybridisation probes, primers  
CC for PCR, for chromosome and gene mapping, for the recombinant production  
CC of protein, or in generation of anti-sense DNA or RNA. The  
CC polynucleotides are useful in diagnostics as expressed sequence tags  
CC (ESTs) for identifying expressed genes or for physical mapping of the  
CC human genome. The proteins may be used as molecular weight markers, or as  
CC nutritional sources or supplements. The proteins may be used to maintain  
CC and expand cell population in a totipotent or pluripotent state  
CC useful for re-engineering damaged or diseased tissues, transplantation,  
CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The  
CC polynucleotides and proteins are useful for preventing, treating or  
CC ameliorating disorders involving aberrant protein expression or  
CC biological activity, e.g. haematopoietic disorders, central/peripheral  
CC nervous system diseases, mechanical and traumatic disorders, non-healing  
CC wounds, immune deficiencies and disorders, infectious diseases caused by  
CC viral, bacterial or fungal infection, autoimmune disorders, allergic  
CC reactions and conditions, coagulation disorders, or cancer. The  
CC polynucleotide sequences of the invention were assembled from ESTs  
CC isolated mainly by sequencing by hybridisation, and in some cases,  
CC sequences obtained from one or more public databases. Note: The sequence  
CC data for this patent did not form part of the printed specification, but  
CC was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 273 AA;  
XX  
Query Match 74.0%; Score 37; DB 5; Length 273;  
Best Local Similarity 55.6%; Pred. No. 1.5e+02;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLWPEIDIV 9  
| | | : : : |  
Db 10 FLWLQDLV 18

## RESULT 10

ABBS4865  
ID ABB54865 standard; protein; 528 AA.

XX ABB54865;

XX 29-AUG-2003 (revised)

DT 16-MAY-2002 (first entry)

XX Lactococcus lactis protein ypsd.

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX Lactococcus lactis; IL1403.

XX FR2807446-A1.

XX 12-OCT-2001.

XX 11-APR-2000; 2000FR-00004630.

XX 11-APR-2000; 2000FR-00004630.

XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.

XX Bolotline A, Sorokine A, Renault P, Ehrlich SD;

XX WPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification or Lactococcus  
lactis and related species.

XX Claim 6; SEQ ID NO 1567; 2504pp; French.

XX The present invention is related to a Lactococcus lactis nucleotide  
sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic  
acid sequence is useful in the detection and/or amplification of nucleic  
acid sequence, particularly to identify Lactococcus lactis or related  
species. The proteins of the invention are useful for the biosynthesis or  
biodegradation of a composition of interest. The invention helps research  
in lactic bacteria, particularly useful in the production of yogurt and  
cheese. Note: The sequence data for this patent is based on equivalent  
CC patent WO20017734 (published 18-OCT-2001) which is available in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences. (Updated on 29-AUG-2003 to  
CC standardise OS field)

XX Sequence 528 AA;

Query Match 74.0%; Score 37; DB 5; Length 528;

Best Local Similarity 66.7%; Pred. No. 3.1e+02; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWPEIDIV 9  
| | | : : : |  
Db 49 FLWLAIIDII 57

## RESULT 11

ABM70700  
ID ABB70700 standard; protein; 587 AA.

XX ABB70700;

XX 20-NOV-2003 (first entry)

XX

DE Photorhabdus luminescens protein sequence #3797.

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
XX detection; food; gene expression; plant; animal; microorganism; toxin;  
XX antibiotic; bioplastic; virulence factor; disease model; plague;  
XX whooping cough.

XX Photorhabdus luminescens.

XX WO200294867-A2.

XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.

XX 07-FEB-2001; 2001FR-00001659.

XX (INSP ) INST PASTEUR.

XX (CNRS ) CNRS CENT NAT RECH SCI.

XX Duchaud E, Taourit S, Glaefer P, Frangeul L, Kunst F, Danchin A;  
XX Buchrieser C;

XX WPI; 2003-148459/14.

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX Claim 2; SEQ ID NO 3797; 1205pp; French.

XX The invention relates to the isolation of genes and their encoded  
XX proteins from Photorhabdus luminescens. The isolated sequences are  
XX sources of probes and primers for detecting the genome of P. luminescens  
XX and related species; to study polymorphisms; for gene analysis and for  
XX detection/amplification of the genes. Antibodies (Ab) raised against the  
XX polypeptides encoded by the genes are used for detection/identification  
XX of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that  
XX carry a gene-containing vector are used to select compounds that  
XX modulate, regulate, induce or inhibit expression of the genes in plants,  
XX animals or microorganisms other than P. luminescens and are able to alter  
XX response or sensitivity to toxins and antibiotics produced by P.  
XX luminescens. Cells transformed to express the genes are useful for  
XX recombinant production of the proteins, particularly toxins and  
XX antibiotics useful as insecticides, bactericides and fungicides. The  
XX genes, proteins, vectors containing the genes and Ab are also useful  
XX therapeutically (to treat microbial infection by bacteria or fungi that  
XX are sensitive to P. luminescens-encoded toxins or antibiotics) and as  
XX bioplasticides. Other uses of the genes and the proteins are as virulence  
XX factors and for identifying targets of human diseases for which P.  
XX luminescens is a model (particularly plague and whooping cough). This  
XX sequence represents one of the isolated P. luminescens proteins

XX Sequence 587 AA;

Query Match 74.0%; Score 37; DB 6; Length 587;

Best Local Similarity 71.4%; Pred. No. 3.5e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFEIDIV 9  
| | | : : : |  
Db 272 WFEIDIV 278

## RESULT 12

ABBS3431  
ID ABB53431 standard; protein; 269 AA.

XX ABB53431;

XX 29-AUG-2003 (revised)

XX 16-MAY-2002 (first entry)

XX Lactococcus lactis protein ypdC.

```

XX  Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
KM  Lactococcus lactis; IL1403.
XX  OS
XX  FR2807446-A1.
XX  PN
XX  12-OCT-2001.
XX  PD
XX  11-APR-2000; 2000FR-00004630.
XX  PF
XX  11-APR-2000; 2000FR-00004630.
XX  PR
XX  11-APR-2000; 2000FR-00004630.
XX  PA
XX  (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX  P1  Bojottine A, Sorokine A, Renault P, Ehrlich SD;
XX  DR  WPI; 2002-043418/06.
XX  PT  New nucleotide sequence useful in the identification or Lactococcus
XX  PT  lactis and related species.
XX  PS  Claim 6; SEQ ID NO 133; 2504pp; French.
XX  CC  The present invention is related to a Lactococcus lactis nucleotide
XX  CC  sequence (ABA90521) and related proteins (AB53300-AB55621). The nucleic
XX  CC  acid sequence is useful in the detection and/or amplification of nucleic
XX  CC  acid sequence, particularly to identify Lactococcus lactis or related
XX  CC  species. The proteins of the invention are useful for the biosynthesis or
XX  CC  biodegradation of a composition of interest. The invention helps research
XX  CC  in lactic bacteria, particularly useful in the production of yogurt and
XX  CC  cheese. Note: The sequence data for this patent is based on equivalent
XX  CC  patent WO200177334 (published 18-OCT-2001) which is available in
XX  CC  electronic format directly from WIPO at
XX  CC  ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
XX  CC  standardise OS field)
XX  SQ  Sequence 269 AA;

Query Match          72.0%; Score 36; DB 5; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 FLWFEI 6
    |||||
Db  155 FLWFEI 160

RESULT 13
ABG01152
ID  ABG01152 standard; protein; 280 AA.
XX
AC  ABG01152;
XX
DT  13-FEB-2002 (first entry)
XX
DE  Novel human diagnostic protein #1143.
XX
KM  Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX  food supplement; medical imaging; diagnostic; genetic disorder.
XX  OS
XX  Homo sapiens.
XX  PN  WO200175067-A2.
XX  PD  11-OCT-2001.
XX  PF  30-MAR-2001; 2001WO-US008631.
XX  PR  31-MAR-2000; 2000US-00540217.
XX  PR  23-AUG-2000; 2000US-00649167.
XX  PA  (HYSE-) HYSEQ INC.

```

```

XX  Drmanac RT, Liu C, Tang YT;
XX  PI  WPI; 2001-639362/73.
XX  DR  N-PSDB; AAS65339.
XX  DR
XX  PT  New isolated polynucleotide and encoded polypeptides, useful in
XX  PT  diagnostics, forensics, gene mapping, identification of mutations
XX  PT  responsible for genetic disorders or other traits and to assess
XX  PT  biodiversity.
XX
XX  PS  Claim 20; SEQ ID NO 31511; 103pp; English.
XX
XX  CC  The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX  CC  sequences. (I) is useful as hybridisation probes, polymerase chain
XX  CC  reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
XX  CC  and in recombinant production of (II). The polynucleotides are also used
XX  CC  in diagnostics as expressed sequence tags for identifying expressed
XX  CC  genes. (I) is useful in gene therapy techniques to restore normal
XX  CC  activity of (II) or to treat disease states involving (II). (II) is
XX  CC  useful for generating antibodies against it, detecting or quantitating a
XX  CC  polypeptide in tissue, as molecular weight markers and as a food
XX  CC  supplement. (II) and its binding partners are useful in medical imaging
XX  CC  of sites expressing (II). (I) and (II) are useful for treating disorders
XX  CC  involving aberrant protein expression or biological activity. The
XX  CC  polypeptide and polynucleotide sequences have applications in
XX  CC  diagnostics, forensics, gene mapping, identification of mutations
XX  CC  responsible for genetic disorders or other traits to assess biodiversity
XX  CC  and to produce other types of data and products dependent on DNA and
XX  CC  amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX  CC  amino acid sequences of the invention. Note: The sequence data for this
XX  CC  patent did not appear in the printed specification, but was obtained in
XX  CC  electronic format directly from WIPO at
XX  CC  ftp.wipo.int/pub/published_pct_sequences
XX  SQ  Sequence 280 AA;

Query Match          72.0%; Score 36; DB 4; Length 280;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY  1 FLWFEDI 8
    |||||
Db  40 FWFEDDV 47

RESULT 14
ABO75288
ID  ABO75288 standard; protein; 373 AA.
XX
AC  ABO75288;
XX
DT  29-JUL-2004 (first entry)
XX
DE  Pseudomonas aeruginosa polypeptide #7463.
XX
KM  Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX  OS
XX  Pseudomonas aeruginosa.
XX  PN  US6551795-B1.
XX  PD  22-APR-2003.
XX  PF  18-FEB-1999; 99US-00252991.
XX  PR  18-FEB-1998; 98US-0074788P.
XX  PR  27-JUL-1998; 98US-0094190P.
XX  PA  (GENO-) GENOME THERAPEUTICS CORP.
XX  PI  Rubenfield MJ, Nolling J, Deloughery C, Bush D;

```

DR WPI; 2003-615309/58.  
DR N-PSDB; ABD08859.  
PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
PS Disclosure; SEQ ID NO 24034; 455bp; English.  
XX  
CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a polypeptide,  
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-  
CC AB084336 represent *P. aeruginosa* polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
SQ Sequence 373 AA;  
Query Match 72.0%; Score 36; DB 7; Length 373;  
Best Local Similarity 55.6%; Pred. No. 3.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
GY 1 FLMFEDIV 9  
Db 307 YLMVEIDLTV 315  
RESULT 15  
ABG01159  
ID ABG01159 standard; protein; 410 AA.  
XX  
AC ABG01159;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #1150.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
FN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001MO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR N-PSDB; AAS65346.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostic, forensic, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX

PS Claim 20; SEQ ID NO 31518; 103bp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC diagnostics, forensic, gene mapping, identification of mutations  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AB050010-AB030377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 410 AA;  
Query Match 72.0%; Score 36; DB 4; Length 410;  
Best Local Similarity 62.5%; Pred. No. 3.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
GY 1 FLMFEDIV 8  
Db 200 FIMFEDDV 207  
RESULT 16  
ABM68497  
ID ABM68497 standard; protein; 591 AA.  
XX  
AC ABM68497;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Photorhabdus luminescens protein sequence #1594.  
XX  
KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
KW detection; food; gene expression; plant; animal; microorganism; toxin;  
KW antibiotic; biopesticide; virulence factor; disease model; plague;  
KW whooping cough.  
XX  
OS Photorhabdus luminescens.  
XX  
FN WO200294867-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 07-FEB-2002; 2002MO-IB003040.  
XX  
PR 07-FEB-2001; 2001FR-00001659.  
XX  
PA (INSP ) INST PASTEUR.  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
PI Duchaud B, Taourit S, Glaser P, Frangeul L, Kunet F, Danchin A;  
XX  
DR WPI; 2003-148459/14.  
XX  
PT Genomic sequence of *Photorhabdus luminescens* and encoded polypeptides,  
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
XX  
PS Claim 2; SEQ ID NO 1594; 1205bp; French.  
XX

CC The invention relates to the isolation of genes and their encoded  
 CC proteins from *Photobacterium luminescens*. The isolated sequences are  
 CC sources of probes and primers for detecting the genome of *P. luminescens*  
 CC and related species; to study polymorphisms; for gene analysis and for  
 CC detection/amplification of the genes. Antibodies (Ab) raised against the  
 CC polypeptides encoded by the genes are used for detection/identification  
 CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that  
 CC carry a gene-containing vector are used to select compounds that  
 CC modulate, regulate, induce or inhibit expression of the genes in plants,  
 CC animals or microorganisms other than *P. luminescens* and are able to alter  
 CC response or sensitivity to toxins and antibiotics produced by *P.*  
 CC *luminescens*. Cells transformed to express the genes are useful for  
 CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically (to treat microbial infection by bacteria or fungi that  
 CC are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which *P.*  
 CC *luminescens* is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated *P. luminescens* proteins

XX Sequence 591 AA;

Query Match 72.0%; Score 36; DB 6; Length 591;  
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LWFEPID 7  
 |||||  
 Db 255 LWFEPID 260

RESULT 17

ABO14689 standard; protein; 652 AA.

XX ABO14689;

DT 25-AUG-2003 (first entry)

XX Novel human protein #62.

Human; NOV; gene therapy; endocrine related disease; diabetes;  
 KW metabolism-related disease; obesity; central nervous system disorder;  
 KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;  
 KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;  
 KW psoriasis; allergy; lupus erythematosus; asthma; cancer;  
 KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;  
 KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;  
 KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;  
 KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;  
 KW stroke; infection.

XX Homo sapiens.

XX WO2003023002-A2.

XX PD 20-MAR-2003.

XX 09-SEP-2002; 2002WO-US028539.

XX 07-SEP-2001; 2001US-0318120P.

XX 07-SEP-2001; 2001US-0318130P.

XX 10-SEP-2001; 2001US-0318430P.

XX 17-SEP-2001; 2001US-0322636P.

XX 17-SEP-2001; 2001US-0322816P.

XX 17-SEP-2001; 2001US-0322817P.

XX 19-SEP-2001; 2001US-0323519P.

XX 20-SEP-2001; 2001US-0323631P.

XX 20-SEP-2001; 2001US-0323636P.

PR 25-SEP-2001; 2001US-0325091P.  
 PR 26-SEP-2001; 2001US-0324990P.  
 PR 17-APR-2002; 2002US-0373212P.  
 PR 06-SEP-2002; 2002US-00236177.

XX (CPR)- CRRAGEN CORP.

PI Spytek KA, Paturajan M, Gorman L, Li L, Anderson DW, Zhong M;  
 PI Getlach VV, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;  
 PI Shinkens RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;  
 PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsbrook JP;  
 PI Lepley DM, Edinger SR, Burgess CE;  
 XX WPI; 2003-313242/30.  
 DR N-PSDB; ACD19382.

PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOV)  
 PT and polynucleotides, useful in gene therapy, e.g. for treating or  
 PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,  
 PT stroke or infections.

PS Claim 1; Page 195-196; 586pp; English.

CC The invention describes a new isolated polypeptide (NOV). The NOV  
 CC polypeptide, nucleic acid and antibody are useful as therapeutics,  
 CC particularly in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease, which includes a pathology associated  
 CC with NOV polypeptide. The DNA encoding the protein is useful in gene  
 CC therapy for treating the disease or condition. In particular, the NOV  
 CC polypeptide or polynucleotide is useful for treating endocrine/  
 CC metabolism-related diseases (e.g. obesity or diabetes), central nervous  
 CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,  
 CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune  
 CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,  
 CC asthma, inflammatory bowel disease, rheumatoid arthritis or  
 CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,  
 CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver  
 CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),  
 CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).  
 CC These are also useful in developing powerful assay system for functional  
 CC analysis of various human disorders, as well as in diagnostic  
 CC applications, and for monitoring the effects of drugs during clinical  
 CC trials. This is the amino acid sequence of a novel human NOV protein

XX Sequence 652 AA;

Query Match 72.0%; Score 36; DB 6; Length 652;  
 Best Local Similarity 75.0%; Pred. No. 5.9e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LWFEPID 8  
 |||||  
 Db 564 LWFEPID 571

RESULT 18

ABO14690 standard; protein; 700 AA.

XX ABO14690;

DT 25-AUG-2003 (first entry)

XX Novel human protein #63.

Human; NOV; gene therapy; endocrine related disease; diabetes;  
 KW metabolism-related disease; obesity; central nervous system disorder;  
 KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;  
 KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;  
 KW psoriasis; allergy; lupus erythematosus; asthma; cancer;  
 KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;  
 KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;  
 KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;

KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;  
stroke; infection.  
XX Homo sapiens.  
XX WO2003023002-A2.  
XX 20-MAR-2003.  
XX 09-SEP-2002; 2002WO-US028539.  
XX 07-SEP-2001; 2001US-0318120P.  
XX 07-SEP-2001; 2001US-0318130P.  
XX 10-SEP-2001; 2001US-0318430P.  
XX 17-SEP-2001; 2001US-0322636P.  
XX 17-SEP-2001; 2001US-0322781P.  
XX 17-SEP-2001; 2001US-0322816P.  
XX 17-SEP-2001; 2001US-0322817P.  
XX 19-SEP-2001; 2001US-0323519P.  
XX 20-SEP-2001; 2001US-0323631P.  
XX 20-SEP-2001; 2001US-0324699P.  
XX 25-SEP-2001; 2001US-0325091P.  
XX 25-SEP-2001; 2001US-0324990P.  
XX 17-APR-2002; 2002US-0373212P.  
XX 06-SEP-2002; 2002US-00236177.  
XX (CURA-) CURAGEN CORP.  
XX Spylek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M,  
XX Gerlach VL, Vernet CM, Ellerman K, Berghs C, Rothenberg ME, Guo X,  
XX Shinkens RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE,  
XX Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigar M, Alsobrook JP,  
XX Lepley DM, Edinger SR, Burgess CE,  
XX WPI; 2003-313242/30.  
XX N-PSDB; ACD19383.  
XX DR  
XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)  
XX and polynucleotides, useful in gene therapy, e.g. for treating or  
XX preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,  
XX stroke or infections.  
XX PT  
XX  
XX PS Claim 1; Page 196; 5866P; English.  
XX  
XX The invention describes a new isolated polypeptide (NOVX). The NOVX  
XX polypeptide, nucleic acid and antibody are useful as therapeutics,  
XX particularly in the manufacture of a medicament for treating a syndrome  
XX associated with a human disease, which includes a pathology associated  
XX with NOVX polypeptide. The DNA encoding the protein is useful in gene  
XX therapy for treating the disease or condition. In particular, the NOVX  
XX polypeptide or polynucleotide is useful for treating endocrine/  
XX metabolism-related diseases (e.g. obesity or diabetes), central nervous  
XX system disorders (e.g. Alzheimer's disease, Parkinson's disease,  
XX epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune  
XX and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,  
XX asthma, inflammatory bowel disease, rheumatoid arthritis or  
XX osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,  
XX prostate or brain cancers, or melanoma), liver diseases (e.g. liver  
XX cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),  
XX haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).  
XX These are also useful in developing powerful assay system for functional  
XX analysis of various human disorders, as well as in diagnostic  
XX applications, and for monitoring the effects of drugs during clinical  
XX trials. This is the amino acid sequence of a novel human NOV protein  
XX  
XX SQ Sequence 700 AA;

Query Match 72.0%; Score 36; DB 6; Length 700;  
Best Local Similarity 75.0%; Pred. No. 6.3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWFEIDI 8

Db 612 FLWFEIDM 619  
|||||  
RESULT 19  
ADD14125  
ID ADD14125 standard; protein; 700 AA.  
XX ADD14125;  
XX  
XX 01-JAN-2004 (first entry)  
XX  
XX Human src biomarker polypeptide SEQ ID NO:314.  
XX  
XX predictor set; protein tyrosine kinase activity modulator;  
XX protein tyrosine kinase pathway; protein tyrosine kinase; cytoskeletal;  
XX gene therapy; drug sensitivity; genetic profile; cancer; human.  
XX  
XX Homo sapiens.  
XX WO2003062395-A2.  
XX 31-JUL-2003.  
XX  
XX 17-JAN-2003; 2003WO-US001981.  
XX  
XX 18-JAN-2002; 2002US-0350061P.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
XX Huang F, Fairchild CR, Lee FY, Shaw P;  
XX WPI; 2003-636735/60.  
XX N-PSDB; ADD14725.  
XX  
XX New polynucleotides and polypeptides for predicting the activity of  
XX PT compounds that interact with protein tyrosine kinases and/or protein  
XX tyrosine kinase pathways.  
XX  
XX PS Claim 10; SEQ ID NO 314; 1399P; English.  
XX  
XX The present invention describes a predictor set comprising a plurality of  
XX polynucleotides or polypeptides whose expression pattern is predictive of  
XX the response of cells to treatment with a compound that modulates protein  
XX tyrosine kinase activity or members of the protein tyrosine kinase  
XX pathway. Also described: (1) predicting whether a compound is capable of  
XX modulating the activity of cells, comprising obtaining a sample of cells,  
XX determining whether the cells express a plurality of markers, and  
XX correlating the expression of the markers to the compound's ability to  
XX modulate the activity of the cells; (2) a plurality of cell lines for  
XX identifying polynucleotides and polypeptides whose expression levels  
XX correlate with compound sensitivity or resistance of cells associated  
XX with a disease state; and (3) identifying polynucleotides and  
XX polypeptides that predict compound sensitivity or resistance of cells  
XX associated with a disease state, comprising subjecting the plurality of  
XX cell lines to one or more compounds, analysing the expression pattern of  
XX a microarray of polynucleotides or polypeptides, and selecting or  
XX polynucleotides or polypeptides that predict the sensitivity or  
XX resistance of cells associated with a disease state by using the  
XX expression pattern of the microarray. The polynucleotides and  
XX polypeptides have cytoskeletal activities, and can be used in gene therapy.  
XX The polynucleotides and polypeptides are useful in predicting the  
XX activity of compounds that interact with protein tyrosine kinases and/or  
XX protein tyrosine kinase pathways. These may be used in determining drug  
XX sensitivity in patients to allow the development of individualized  
XX genetic profiles which aid in treating diseases and disorders (e.g.  
XX cancer) based on patient response at a molecular level. The present  
XX sequence is used in the exemplification of the present invention.  
XX  
XX SQ Sequence 700 AA;

Query Match 72.0%; Score 36; DB 7; Length 700;  
Best Local Similarity 75.0%; Pred. No. 6.3e+02;



Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 FLWFEIDI 8  
Db 612 FLWFEIDM 619

RESULT 20  
ADJ70236 standard; protein; 700 AA.

ADJ70236;

06-MAY-2004 (first entry)

Human heat mitochondrial protein as a therapeutic target SegID2042.

microchondrial; human; screening assay; diabetes mellitus;

Huntington's disease; osteoarthritis;

Leber's hereditary optic neuropathy; LHON;

mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;

osteopathic; ophthalmological; cyostatic.

Homo sapiens.

MO2003087768-A2.

23-OCT-2003.

04-APR-2003; 2003WO-US010870.

12-APR-2002; 2002US-0372843P.

17-JUN-2002; 2002US-038987P.

20-SEP-2002; 2002US-0412418P.

(MITO-) MITOKOR.

(BUCK-) BUCK INSR AGE RES.

Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;

WPI; 2003-845369/78.

Identifying a mitochondrial target for drug screening assays and for

treating diseases associated with altered mitochondrial function,

comprising detecting a modified polypeptide in a sample and correlating

with the disease.

Claim 1; SEQ ID NO 2042; 180pp; English.

This invention relates to novel mitochondrial targets that can be used

for therapeutic intervention in treating a disease associated with

altered mitochondrial function. Specifically, it refers to a method for

identifying proteins of the human heart mitochondrial proteome that are

useful for drug screening assays, as well as therapeutic targets. The

present invention describes a method for identifying such proteins that

can be used in the treatment of various diseases associated with altered

mitochondrial function including diabetes mellitus, Huntington's disease,

osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial

encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy

ragged red fibre syndrome (MERRF) or cancer. Accordingly, these

compositions have neuroprotective, nootropic, antidiabetic,

anticonvulsant, antiarthritic, osteopathic, ophthalmological and

cyostatic activities. This polypeptide sequence is a human heart

mitochondrial protein of the invention.

Sequence 700 AA;

Query Match 72.0%; Score 36; DB 7; Length 700;

Best Local Similarity 75.0%; Pred. No. 6.3e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLWFEIDI 8  
Db 612 FLWFEIDM 619

RESULT 21  
ABG01151 standard; protein; 1097 AA.

ABG01151;

13-FEB-2002 (first entry)

Novel human diagnostic protein #1142.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

MO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSB-) HYSBQ INC.

Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

N-PSDB; AAS65338.

New isolated polynucleotide and encoded polypeptides, useful in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess

biodiversity.

Claim 20; SEQ ID NO 31510; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II)

sequences. (I) is useful as hybridisation probes, polymerase chain

reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

and in recombinant production of (II). The polynucleotides are also used

in diagnostics as expressed sequence tags for identifying expressed

genes. (I) is useful in gene therapy techniques to restore normal

activity of (II) or to treat disease states involving (II). (II) is

useful for generating antibodies against it, detecting or quantitating a

polypeptide in tissue, as molecular weight markers and as a food

supplement. (II) and its binding partners are useful in medical imaging

of sites expressing (II). (I) and (II) are useful for treating disorders

involving aberrant protein expression or biological activity. The

polynucleotide and polynucleotide sequences have applications in

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits to assess biodiversity

and to produce other types of data and products dependent on DNA and

amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

amino acid sequences of the invention. Note: The sequence data for this

patent did not appear in the printed specification, but was obtained in

electronic format directly from WIPO at

ftp.wipo.int/pub/published\_pct\_sequences

Sequence 1097 AA;

Query Match 72.0%; Score 36; DB 4; Length 1097;

Best Local Similarity 62.5%; Pred. No. 1e+03;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 270 FIMFEDDV 277

# RESULT 22

AD117021

AD117021 standard; protein; 217 AA.

AC AD117021;

DT 15-APR-2004 (first entry)

DE Human NOVX protein homologue SegID 557.

XX human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;

XX inflammation; autoimmune disorder; allergy; blood disorder;

KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;

KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;

KW Alzheimer's disease; infection; str.

OS Homo sapiens.

XX WO200268649-A2.

XX 06-SEP-2002.

XX 31-JAN-2002; 2002WO-US002785.

XX 31-JAN-2001; 2001US-0265395P.

XX 31-JAN-2001; 2001US-0265412P.

XX 31-JAN-2001; 2001US-0265514P.

XX 02-FEB-2001; 2001US-0266406P.

XX 05-FEB-2001; 2001US-0266767P.

XX 07-FEB-2001; 2001US-0266755P.

XX 08-FEB-2001; 2001US-0267057P.

XX 09-FEB-2001; 2001US-0267823P.

XX 15-FEB-2001; 2001US-0268974P.

XX 26-FEB-2001; 2001US-0271664P.

XX 27-FEB-2001; 2001US-0271839P.

XX 02-MAR-2001; 2001US-0272788P.

XX 14-MAR-2001; 2001US-0275925P.

XX 14-MAR-2001; 2001US-0275947P.

XX 14-MAR-2001; 2001US-0275950P.

XX 15-MAR-2001; 2001US-0276448P.

XX 15-MAR-2001; 2001US-0276450P.

XX 16-MAR-2001; 2001US-0276397P.

XX 16-MAR-2001; 2001US-0276768P.

XX 20-MAR-2001; 2001US-0278652P.

XX 26-MAR-2001; 2001US-0278775P.

XX 29-MAR-2001; 2001US-0279882P.

XX 29-MAR-2001; 2001US-0279884P.

PR 28-AUG-2001; 2001US-0315470P.

PR 31-AUG-2001; 2001US-0316447P.

PR 07-SEP-2001; 2001US-0318115P.

PR 07-SEP-2001; 2001US-0318118P.

PR 12-SEP-2001; 2001US-0318740P.

PR 19-SEP-2001; 2001US-0323379P.

PR 18-OCT-2001; 2001US-0330245P.

PR 18-OCT-2001; 2001US-0330308P.

PR 14-NOV-2001; 2001US-0332701P.

XX (CURA-) CURAGEN CORP.

XX Tchernev VT, Spytek KA, Zernhusen BD, Patcurajan M, Shimkets RA;

PI Li L, Ganggoli EA, Padigaru M, Anderson DM, Rastelli L, Miller CE;

PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;

PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;

XX WPI, 2002-706998/76.

XX New NOVX polypeptides and nucleic acids, useful for preventing or

PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,

PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX Disclosure, SEQ ID NO 557; 1498bp; English.

XX This invention relates to a novel nucleic acid, and encoded polypeptides

CC thereof, which have properties related to the stimulation of biochemical

CC or physiological responses in a cell, tissue, organ or organism.

CC Specifically, it refers to the use of biologically active fragments for

CC diagnostic and prognostic assays and furthermore in the treatment of

CC diverse pathological conditions. The present invention describes novel

CC human and murine NOVX proteins, as well as methods to modulate their

CC expression using antisense oligos, ribozymes and peptide nucleic acids.

CC The NOVX polypeptides, polynucleotides and antibodies are useful in

CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,

CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in

CC treating or preventing diseases such as inflammation, autoimmune

CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome

CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,

CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy

CC and epilepsy. Accordingly, these molecules have many activities including

CC cytoskeletal, cardiant, anti-inflammatory, immunosuppressive, antiallergic,

CC hemostatic, anti-HIV, antidiabetic, antilactosclerotic, anorectic,

CC antiaesthetic, nephroprotective, antiarthritic, hepatotropic,

CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,

CC relaxant and anticoagulant. In addition, they are useful in screening

CC assays to identify small molecules that modulate or inhibit, for example,

CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also

CC used as in chromosome mapping, tissue typing, preventive medicine and

CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein

CC of the invention.

XX Sequence 217 AA;

XX Query Match 70.0%; Score 35; DB 5; Length 217;

XX Best Local Similarity 77.8%; Pred. No. 2.7e+02;

XX Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFELDIV 9

Db 86 FLWFELDIV 94

## RESULT 23

ID ADC86399 standard; protein; 232 AA.

AC ADC86399;

DT 01-JAN-2004 (first entry)

XX Human GPCR protein SEQ ID NO:852.

XX human; GPCR; guanosine triphosphate-binding protein coupled receptor;  
 KW gene therapy.  
 XX Homo sapiens.  
 XX EPI270724-A2.  
 XX 02-JAN-2003.  
 XX PD  
 XX 18-JUN-2002; 2002EP-00013517.  
 XX 18-JUN-2001; 2001JP-00246789.  
 XX PR  
 XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
 XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.  
 XX PA  
 XX Suwa M, Asai K, Akiyama Y, Aburatani H;  
 XX WPI; 2003-315783/31.  
 XX DR N-PSDB; ADC86398.  
 XX PT New polynucleotide, useful for preparing a composition for treating a  
 PT patient in need of increased or suppressed activity or expression of the  
 PT guanosine triphosphate-binding protein coupled receptor.  
 PS Claim 2; SEQ ID NO 852; 28pp; English.  
 XX  
 XX The invention relates to a novel polynucleotide encoding a guanosine  
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of  
 CC the invention may have a use in gene therapy. The polynucleotide and  
 CC polypeptide are useful for preparing a composition for treating a patient  
 CC in need of increased or suppressed activity or expression of the  
 CC guanosine triphosphate-binding protein coupled receptor. The protein  
 CC sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.  
 CC  
 SQ Sequence 232 AA;  
 Query Match 70.0%; Score 35; DB 7; Length 232;  
 Best Local Similarity 77.8%; Pred. No. 2.9e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 FLMFEIDIV 9  
 DB 168 FLMFLIPIV 176  
 RESULT 24  
 AAG71459  
 ID AAG71459 standard; protein; 307 AA.  
 XX  
 XX AAG71459;  
 AC  
 XX 31-JUL-2001 (first entry)  
 DT  
 XX Human olfactory receptor polypeptide, SEQ ID NO: 1140.  
 DE  
 XX Human; olfactory receptor; OR; primary scent determination;  
 KW secondary scent determination; polypeptide library; odour receptor;  
 KW scent profile; scent fingerprint; scent representation.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200127158-A2.  
 XX PN  
 XX 19-APR-2001.  
 XX PD  
 XX 06-OCT-2000; 2000WO-US027582.  
 XX PF  
 XX 08-OCT-1999; 99US-0158615P.  
 XX PR 24-FEB-2000; 2000US-0184809P.  
 XX PS  
 XX (DIGI-) DIGISCENTS.

PA (YEDA ) YEDA RES & DEV CO LTD.  
 XX  
 XX Piellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
 XX WPI; 2001-290713/30.  
 XX DR  
 XX New polynucleotides which encode polypeptides involved in olfactory  
 PT sensation for identifying olfactory agonists and antagonists.  
 PT  
 XX Claim 11; Page 661-662; 1857pp; English.  
 PS  
 XX The present sequence is an olfactory receptor which is encoded by one of  
 CC a number of novel polynucleotides. The polynucleotides can be used in  
 CC screening for olfactory agonists and antagonists. The methods allow for  
 CC the determination of primary scents and the identification of the odour  
 CC receptors used to detect these primary scents. The methods also enable  
 CC determination of secondary scents and the identification of combinations  
 CC of odour receptors that are involved in detecting such secondary scents.  
 CC This enables the construction of a scent representation (also called a  
 CC scent fingerprint or scent profile), which may be used to re-create and  
 CC edit scents. Libraries of olfactory receptors are useful for determining  
 CC the interaction pattern of a composition with the receptors, and can be  
 CC used for determining differences in the olfactory faculties of different  
 CC individuals  
 CC  
 SQ Sequence 307 AA;  
 Query Match 70.0%; Score 35; DB 4; Length 307;  
 Best Local Similarity 77.8%; Pred. No. 4e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 FLMFEIDIV 9  
 DB 146 FLMFLIPIV 154  
 RESULT 25  
 AAG72215  
 ID AAG72215 standard; protein; 315 AA.  
 XX  
 XX AAG72215;  
 AC  
 XX 30-JUL-2001 (first entry)  
 DT  
 XX Human olfactory receptor polypeptide, SEQ ID NO: 1896.  
 DE  
 XX Human; olfactory receptor; OR; primary scent determination;  
 KW secondary scent determination; polypeptide library; odour receptor;  
 KW scent profile; scent fingerprint; scent representation.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200127158-A2.  
 XX PN  
 XX 19-APR-2001.  
 XX PD  
 XX 06-OCT-2000; 2000WO-US027582.  
 XX PF  
 XX 08-OCT-1999; 99US-0158615P.  
 XX PR 24-FEB-2000; 2000US-0184809P.  
 XX PS  
 XX (DIGI-) DIGISCENTS.  
 XX (YEDA ) YEDA RES & DEV CO LTD.  
 XX  
 XX Piellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
 XX WPI; 2001-290713/30.  
 XX DR  
 XX New polynucleotides which encode polypeptides involved in olfactory  
 PT sensation for identifying olfactory agonists and antagonists.  
 PT  
 XX Claim 11; Page 1258-1259; 1857pp; English.  
 PS

CC The present sequence is an olfactory receptor which is encoded by one of  
 CC a number of novel polynucleotides. The polynucleotides can be used in  
 CC screening for olfactory agonists and antagonists. The methods allow for  
 CC the determination of primary scents and the identification of the odour  
 CC receptors used to detect these primary scents. The methods also enable  
 CC determination of secondary scents and the identification of combinations  
 CC of odour receptors that are involved in detecting such secondary scents.  
 CC This enables the construction of a scent representation (also called a  
 CC scent fingerprint or scent profile), which may be used to re-create and  
 CC edit scents. Libraries of olfactory receptors are useful for determining  
 CC the interaction pattern of a composition with the receptors, and can be  
 CC used for determining differences in the olfactory faculties of different  
 CC individuals

CC Sequence 315 AA:

SO Query Match 70.0%; Score 35; DB 4; Length 315;  
 Best Local Similarity 77.8%; Pred. No. 4.1e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
 |||||  
 Db 157 FLWFLIPIV 165

RESULT 26

ABG79345  
 ID ABG79345 standard; protein; 315 AA.

AC ABG79345;

DT 15-NOV-2002 (first entry)

DE Human GPCR7 protein.

KW G protein coupled receptor related protein; human; GPCR; cardiomyopathy;  
 KW atherosclerosis; diabetes; cancer; stroke; Von Hippel-Lindau syndrome;  
 KW Alzheimer's disease; tuberosus sclerosis; hypercalcaemia;  
 KW Parkinson's disease; Huntington's disease; cerebral palsy; epilepsy;  
 KW Lesch-Nyhan syndrome; multiple sclerosis; ataxia-telangiectasia;  
 KW leukodystrophy; addiction; anxiety; depression; pain; obesity;  
 KW Crohn's disease; osteoporosis; haemophilia; asthma;  
 KW inflammatory bowel disease; infertility; hypertension; scleroderma;  
 KW arthritis; human immunodeficiency virus; autoimmune disease; HIV;  
 KW infection; graft-versus-host disease.

OS Homo sapiens.

PN WO200264793-A2.

PD 22-AUG-2002.

PF 03-JAN-2002; 2002WO-US000056.

PR 03-JAN-2001; 2001US-0259552P.

PR 09-JAN-2001; 2001US-0260544P.

PR 20-MAR-2001; 2001US-0277405P.

PA (CURA-) CURAGEN CORP.

PI Casman SJ, Edinger SR, Ellerman K, Smithson G, Kekuda R;

PI Muralidhara P;

DR WPI; 2002-643487/69.

XX N-PSDB; ABS64714.

PT New isolated G protein coupled receptor polypeptides and polynucleotides,  
 PT useful in gene therapy, particularly for treating or preventing  
 PT cardiomyopathy, atherosclerosis, diabetes, multiple sclerosis, Crohn's  
 PT disease or cancer.

PS Claim 1; Page 50; 252pp; English.

CC This invention relates to a new isolated G-protein coupled receptor  
 CC (GPCRX) polypeptide sequence and the cDNA encoding it. The GPCR  
 CC polypeptide, GPCRX nucleic acid and an antibody specific to the protein  
 CC are useful for treating, preventing or alleviating a GPCR-associated  
 CC disorder or a pathological state in a subject, particularly a human. In  
 CC particular, the disorder is cardiomyopathy, atherosclerosis, diabetes, or  
 CC a disorder related to cell signal processing and metabolic pathway  
 CC modulation. The GPCRX polypeptide and nucleic acid are also useful for  
 CC diagnosing the presence of or predisposition to a disease associated with  
 CC altered levels of GPCRX, particularly cancer. The GPCRX nucleic acid and  
 CC polypeptide are especially useful in the manufacture of a medicament for  
 CC therapeutic or prophylactic applications for disorders associated with  
 CC aberrant GPCRX expression or activity, e.g. Von Hippel-Lindau syndrome,  
 CC Alzheimer's disease, stroke, tuberosus sclerosis, hypercalcaemia,  
 CC Parkinson's disease, Huntington's disease, cerebral palsy, epilepsy,  
 CC Lesch-Nyhan syndrome, multiple sclerosis, ataxia-telangiectasia,  
 CC leukodystrophies, addiction, anxiety, depression, pain, obesity, Crohn's  
 CC disease, osteoporosis, inflammatory bowel disease, infertility,  
 CC hypertension, scleroderma, haemophilia, asthma, arthritis, human  
 CC immunodeficiency virus; autoimmune disease; HIV, viral, fungal, bacterial  
 CC or protozoal infections, or graft-versus -host disease. The DNA encoding  
 CC the protein is useful in gene therapy for treating the above conditions.  
 CC The polypeptides can be used as immunogens to produce antibodies and as  
 CC vaccines. The nucleic acids are further used as hybridisation probes, in  
 CC chromosome mapping, tissue typing, preventive medicine, and  
 CC pharmacogenomics. These are also useful in developing powerful assay  
 CC system for functional analysis of various human disorders, as well as in  
 CC diagnostic applications. The present sequence represents a human G  
 CC protein coupled receptor related protein (GPCR) of the invention

CC Sequence 315 AA:

SO Query Match 70.0%; Score 35; DB 5; Length 315;  
 Best Local Similarity 77.8%; Pred. No. 4.1e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
 |||||  
 Db 157 FLWFLIPIV 165

RESULT 27

AD116656  
 ID AD116656 standard; protein; 315 AA.

AC AD116656;

DT 15-APR-2004 (first entry)

DE Human NOXV protein to treat human pathological conditions SeqID192.

KW human; NOXV; cardiomyopathy; atherosclerosis; cancer; diabetes;  
 KW inflammation; autoimmune disorder; allergy; blood disorder;  
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
 KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;  
 KW Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;  
 KW cytosarctic; cardiac; antiinflammatory; immunosuppressive; antiallergic;  
 KW haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic; anorectic;  
 KW antiaesthetic; nephrotropic; antiarthritic; hepatotropic;  
 KW neuroprotective; nootropic; antibacterial; virucide; antiparasitic;  
 KW relaxant; anticonvulsant; neurogenesis; wound healing; angiogenesis;  
 KW chromosome mapping; tissue typing; pharmacogenomic.

OS Homo sapiens.

PN WO200268649-A2.

PD 06-SEP-2002.

PF 31-JAN-2002; 2002WO-US002785.

PR 31-JAN-2001; 2001US-0265395P.

PR 31-JAN-2001; 2001US-0265412P.

PR 31-JAN-2001; 2001US-0265511P.  
PR 31-JAN-2001; 2001US-0265517P.  
PR 02-FEB-2001; 2001US-0266406P.  
PR 05-FEB-2001; 2001US-0266767P.  
PR 07-FEB-2001; 2001US-0266975P.  
PR 07-FEB-2001; 2001US-0267057P.  
PR 08-FEB-2001; 2001US-0267453P.  
PR 09-FEB-2001; 2001US-0267823P.  
PR 15-FEB-2001; 2001US-0268974P.  
PR 26-FEB-2001; 2001US-0271664P.  
PR 27-FEB-2001; 2001US-0271839P.  
PR 27-FEB-2001; 2001US-0271855P.  
PR 02-MAR-2001; 2001US-0272788P.  
PR 02-MAR-2001; 2001US-0273046P.  
PR 14-MAR-2001; 2001US-0275925P.  
PR 14-MAR-2001; 2001US-0275947P.  
PR 14-MAR-2001; 2001US-0275950P.  
PR 15-MAR-2001; 2001US-0275989P.  
PR 15-MAR-2001; 2001US-0276448P.  
PR 15-MAR-2001; 2001US-0276450P.  
PR 16-MAR-2001; 2001US-0276397P.  
PR 16-MAR-2001; 2001US-0276768P.  
PR 20-MAR-2001; 2001US-0278652P.  
PR 26-MAR-2001; 2001US-0278775P.  
PR 26-MAR-2001; 2001US-0278778P.  
PR 29-MAR-2001; 2001US-0279882P.  
PR 29-MAR-2001; 2001US-0279884P.  
PR 30-MAR-2001; 2001US-0280147P.  
PR 11-APR-2001; 2001US-0282993P.  
PR 11-APR-2001; 2001US-0283083P.  
PR 20-APR-2001; 2001US-0285133P.  
PR 23-APR-2001; 2001US-0285749P.  
PR 03-MAY-2001; 2001US-0288327P.  
PR 03-MAY-2001; 2001US-0288504P.  
PR 29-MAY-2001; 2001US-0294047P.  
PR 30-MAY-2001; 2001US-0294473P.  
PR 08-JUN-2001; 2001US-0296964P.  
PR 18-JUN-2001; 2001US-0298959P.  
PR 19-JUN-2001; 2001US-0299324P.  
PR 13-AUG-2001; 2001US-0312020P.  
PR 16-AUG-2001; 2001US-0312889P.  
PR 16-AUG-2001; 2001US-0312908P.  
PR 21-AUG-2001; 2001US-031390P.  
PR 21-AUG-2001; 2001US-031390P.  
PR 28-AUG-2001; 2001US-0315470P.  
PR 31-AUG-2001; 2001US-0316447P.  
PR 07-SEP-2001; 2001US-0318115P.  
PR 07-SEP-2001; 2001US-0318118P.  
PR 12-SEP-2001; 2001US-0318740P.  
PR 19-SEP-2001; 2001US-032379P.  
PR 18-OCT-2001; 2001US-0330245P.  
PR 18-OCT-2001; 2001US-0330308P.  
PR 14-NOV-2001; 2001US-0332701P.  
XX  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
XX Tchernev VT, Spytek KA, Zernhusen BD, Paturajan M, Shimkete RA;  
PI Li L, Ganggoli EA, Padigaru M, Anderson DW, Pastelli L, Miller CE;  
PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;  
PI Putrak K, Grose WM, Alabrook JP, Lepley DM, Rieger DK, Burgess CE;  
XX  
XX WPI; 2002-706998/76.  
DR N-PSDB; ADI16655.  
XX  
XX  
PT New NOVX polypeptides and nucleic acids, useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or  
PT pharmacogenomics.  
XX  
XX  
PS Claim 1; SEQ ID NO 192; 1498bp; English.  
XX  
XX  
CC This invention relates to a novel nucleic acids, and encoded polypeptides  
CC thereof, which have properties related to the stimulation of biochemical  
CC or physiological responses in a cell, tissue, organ or organism.

CC Specifically, it refers to the use of biologically active fragments for  
CC diagnostic and prognostic assays and furthermore in the treatment of  
CC diverse pathological conditions. The present invention describes novel  
CC human and murine NOVX proteins, as well as methods to modulate their  
CC expression using antisense oligos, ribozymes and peptide nucleic acids.  
CC The NOVX polypeptides, polynucleotides and antibodies are useful in  
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,  
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in  
CC treating or preventing diseases such as inflammation, autoimmune  
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome  
CC (AIDS), obesity, asthma, immunoglobulin (Ig) A nephropathy, cirrhosis,  
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy  
CC and epilepsy. Accordingly, these molecules have many activities including  
CC cyostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,  
CC haemostatic, anti-HIV, antidiabetic, antiatherosclerotic, anorectic,  
CC antiaesthetic, nephrotropic, antiarthritic, hepatotropic,  
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,  
CC relaxant and anticonvulsant. In addition, they are useful in screening  
CC assays to identify small molecules that modulate or inhibit, for example,  
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also  
CC used as in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. This polypeptide is a human NOVX protein of the  
CC invention.  
XX  
XX  
SQ Sequence 315 AA;  
Query Match 70.0%; Score 35; DB 5; Length 315;  
Best Local Similarity 77.8%; Pred. No. 4.1e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 FLWFEIDIV 9  
Db 157 FLWFLPIIV 165  
RESULT 28  
ID ABR01643 standard; protein; 315 AA.  
XX  
XX ABR01643;  
AC  
XX  
DT 17-APR-2003 (first entry)  
XX  
XX Human G protein coupled receptor SEQ ID 146.  
XX  
XX Human: anorectic; antidiabetic; hypotensive; receptor;  
XX antiparkinsonian; nootropic; neuroprotective; tranquilizer;  
XX antineumatic; antiinflammatory; osteopathic; cardiant; neuroleptic;  
XX antiarthritic; gene therapy; olfactory G protein-coupled receptor; GPCR;  
XX infection; obesity; diabetes; hypertension; malnutrition;  
XX Parkinson's disease; Alzheimer's disease; Korsakoff's psychosis; anxiety;  
XX rheumatoid arthritis; chronic obstructive pulmonary disease;  
XX osteoporosis; asthma; myocardial infarction; schizophrenia;  
XX osteoarthritis.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2003000735-A2.  
PN  
XX  
XX 03-JAN-2003.  
PD  
XX  
XX 24-JUN-2002; 2002MO-IB002481.  
PF  
XX  
XX 26-JUN-2001; 2001US-0301095P.  
PR 06-NOV-2001; 2001US-0332758P.  
XX  
XX (DECO-) DECODE GENETICS BHF.  
PA  
XX  
XX Martinez RAM, Sigurdson GT;  
PI  
XX  
XX WPI; 2003-175284/17.  
DR N-PSDB; ABZ77944.  
XX

PT New olfactory G protein-coupled receptor gene nucleic acid and  
PT polypeptide, useful for diagnosing or treating a disease or condition  
PT associated with GPCR, e.g. obesity, diabetes, hypertension, malnutrition  
PT or Alzheimer's disease.

XX Claim 9; Page 90; 383pp; English.

CC The present invention relates to novel human olfactory G protein-coupled  
CC receptors (GPCR) and their coding sequences (AB277872-AB277986 and  
CC ABR01571-ABR01685). The GPCRs and coding sequences are useful for  
CC diagnosing or treating a disease or condition associated with GPCR, e.g.  
CC infections, obesity, diabetes, hypertension, malnutrition, Parkinson's  
CC disease, Alzheimer's disease, Korsakoff's psychosis, anxiety, rheumatoid  
CC arthritis, chronic obstructive pulmonary disease, osteoporosis, asthma,  
CC myocardial infarction, schizophrenia, or osteoarthritis

XX Sequence 315 AA;

Query Match Best Local Similarity 77.8%; Score 35; DB 6; Length 315;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLNFEIDIV 9  
Db 154 FLNFEIDIV 162

RESULT 29

ADN42310 ADN42310 standard; protein; 315 AA.

AC ADN42310;

DT 17-JUN-2004 (first entry)

XX Human novel protein NOV 51.

XX Human; NOV; cancer; diabetes; cardiomyopathy; atherosclerosis.

OS Homo sapiens.

PN US2004033493-A1.

PD 19-FEB-2004.

PF 31-JAN-2002; 2002US-00072012.

XX 31-JAN-2001; 2001US-0265395P.

XX 31-JAN-2001; 2001US-0265412P.

XX 31-JAN-2001; 2001US-0265514P.

XX 31-JAN-2001; 2001US-0265517P.

XX 02-FEB-2001; 2001US-0266406P.

XX 05-FEB-2001; 2001US-0266767P.

XX 07-FEB-2001; 2001US-0266975P.

XX 08-FEB-2001; 2001US-0267057P.

XX 09-FEB-2001; 2001US-0267823P.

XX 15-FEB-2001; 2001US-0268974P.

XX 26-FEB-2001; 2001US-0271864P.

XX 27-FEB-2001; 2001US-0271855P.

XX 02-MAR-2001; 2001US-0272788P.

XX 02-MAR-2001; 2001US-0273046P.

XX 14-MAR-2001; 2001US-0275925P.

XX 14-MAR-2001; 2001US-0275947P.

XX 14-MAR-2001; 2001US-0275950P.

XX 15-MAR-2001; 2001US-0276448P.

XX 15-MAR-2001; 2001US-0276450P.

XX 16-MAR-2001; 2001US-0276397P.

XX 16-MAR-2001; 2001US-0276768P.

XX 20-MAR-2001; 2001US-0278652P.

XX 26-MAR-2001; 2001US-0278775P.

PR 26-MAR-2001; 2001US-0278778P.

PR 29-MAR-2001; 2001US-0279882P.

PR 29-MAR-2001; 2001US-0279884P.

PR 30-MAR-2001; 2001US-0280147P.

PR 11-APR-2001; 2001US-0282929P.

PR 11-APR-2001; 2001US-0283083P.

PR 20-APR-2001; 2001US-0285133P.

PR 23-APR-2001; 2001US-0285749P.

PR 03-MAY-2001; 2001US-0288327P.

PR 03-MAY-2001; 2001US-0288504P.

PR 29-MAY-2001; 2001US-0294047P.

PR 30-MAY-2001; 2001US-0294473P.

PR 08-JUN-2001; 2001US-0296564P.

PR 18-JUN-2001; 2001US-0298959P.

PR 19-JUN-2001; 2001US-0299324P.

PR 13-AUG-2001; 2001US-0312020P.

PR 16-AUG-2001; 2001US-0312889P.

PR 16-AUG-2001; 2001US-0312908P.

PR 21-AUG-2001; 2001US-031330P.

PR 28-AUG-2001; 2001US-0315470P.

PR 31-AUG-2001; 2001US-0316447P.

PR 07-SEP-2001; 2001US-0318115P.

PR 07-SEP-2001; 2001US-0318118P.

PR 12-SEP-2001; 2001US-0318740P.

PR 19-SEP-2001; 2001US-0323379P.

PR 18-OCT-2001; 2001US-0330245P.

PR 18-OCT-2001; 2001US-0330308P.

PR 14-NOV-2001; 2001US-0332701P.

XX (TCHE/) TCHERNY V T.

XX (SPYT/) SPYTEK K A.

XX (ZERR/) ZERRHUSEN B D.

XX (PAT/) PATURAJAN M.

XX (SHIM/) SHIMKETS R A.

XX (LILL/) LI L.

XX (GANG/) GANGOLI E A.

XX (PAD/) PADIGARU M.

XX (ANDE/) ANDERSON D W.

XX (RAST/) RASTELLI L.

XX (MILL/) MILLER C E.

XX (GERL/) GERLACH V.

XX (TAUP/) TAUPIER R J.

XX (GUSE/) GUSEV V Y.

XX (COLM/) COLMAN S D.

XX (WOLE/) WOLENC A R.

XX (PENNA/) PENNA C E A.

XX (FURT/) FURTAK K.

XX (GROS/) GROSSE W M.

XX (ALSO/) ALSOBROOK J P.

XX (LEPL/) LEPLEY D M.

XX (RIEG/) RIEGER D K.

XX (BURG/) BURGESS C E.

XX Tchermey VT, Spytak KA, Zerrhuse B, Shinkets RA;

XX Li L, Gangoli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;

XX Gusev V, Taupier RJ, Colman SD, Wolenc AR, Penna CE, Furtak K, Grosse WM, Alsbrook JP, Lepley DM, Rieger DK, Burgess CE;

XX MPI; 2004-180039/17.

XX N-PSDB; ADN42309.

XX Isolated NOVX polypeptides and polynucleotides, useful for preventing

XX diagnosing and/or treating cancer, diabetes, cardiomyopathy and

XX atherosclerosis.

XX Claim 1; SEQ ID NO 192; 1309pp; English.

XX The invention relates isolated 162 NOVX polypeptides (NOV1-NOV99,

XX including splice variants) and the nucleic acids (NA) that encode them.

XX Also included are the mature NOVX proteins (and their encoding

XX polynucleotides), a vector comprising NOVX NA, a cell comprising the

XX vector, an antibody that binds immunospecifically to NOVX, determining

XX the presence or amount of NOVX in a sample, determining the presence or

CC amount of NOVX NA in a sample, identifying an agent that binds to NOVX,  
CC modulating the activity of NOVX, treating or preventing a NOVX-associated  
CC disorder, determining the presence of or predisposition to a disease  
CC associated with altered levels of NOVX and treating a pathological state  
CC in a mammal comprising administering a polypeptide which is at least 95%  
CC identical to NOVX (or fragment). NOVX and NA may be used in the  
CC prevention, treatment and diagnosis of diseases associated with  
CC inappropriate expression and activity of NOVX (e.g. cancer, diabetes,  
CC cardiomyopathy and/or atherosclerosis). The anti-NOVX antibodies and  
CC antagonists may also be used to down regulate expression and activity of  
CC NOVX. The anti-NOVX antibodies may also be used as diagnostic agents for  
CC detecting the presence of NOVX in samples (e.g. by enzyme linked in this  
CC immunosorbent assay (ELISA). The agents and methods may be used in this  
CC way to prevent, diagnose and treat cancer, diabetes, cardiomyopathy  
CC and/or atherosclerosis. The present sequence represents a NOVX protein.  
CC  
SQ Sequence 315 AA;  
Query Match 70.0%; Score 35; DB 8; Length 315;  
Best Local Similarity 77.8%; Pred. No. 4.1e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 FLWFEIDIV 9  
Db 157 FLWFLPIIV 165  
RESULT 30  
ABP95915  
ID ABP95915 standard; protein; 326 AA.  
XX  
AC ABP95915;  
XX  
DT 06-MAR-2003 (first entry)  
XX  
DE Human GPCR polypeptide SEQ ID NO 640.  
XX  
KW Human; GPCR; G protein coupled receptor; signal transduction; olfactory;  
KW drug development; gustatory; taste; fragrance; receptor.  
XX  
OS Homo sapiens.  
XX  
PN WO200216548-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 30-JUL-2001; 2001WO-IB001446.  
XX  
PR 04-AUG-2000; 2000JP-00237818.  
PR 13-FEB-2001; 2001JP-00034434.  
XX  
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
PI Haga T, Takeda S, Mitaku S;  
XX  
PI WPI; 2002-304118/34.  
DR N-PSDB; ABZ43189.  
XX  
PT Database global search for G protein-coupled receptors, proteins and  
PT encoded genes for studying in vivo signal transduction mechanism and  
PT identifying targets for drug development.  
XX  
PS Claim 10; SEQ ID NO 640; 97bp + Sequence Listing; Japanese.  
XX  
XX The invention relates to a method for screening G protein-coupled  
CC receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-  
CC ABP95942) by extracting open-reading frames containing 6-8 transmembrane  
CC domains with 250-1000 amino acid residues to give a gene homologous with  
CC a known GPCR gene. The receptor proteins and encoded genes are useful for  
CC studying in vivo signal transduction mechanism and identifying targets  
CC for drug development e.g. based on olfactory and gustatory receptors in  
CC form of agonists and antagonists by screening intrinsic and extrinsic  
CC ligands as bitter taste inhibitors, taste enhancers and fragrance

CC improve. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 326 AA;  
Query Match 70.0%; Score 35; DB 5; Length 326;  
Best Local Similarity 77.8%; Pred. No. 4.2e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 FLWFEIDIV 9  
Db 168 FLWFLPIIV 176  
RESULT 31  
AAU95516  
ID AAU95516 standard; protein; 326 AA.  
XX  
AC AAU95516;  
XX  
DT 02-JUL-2002 (first entry)  
XX  
DE Human olfactory and pheromone G protein-coupled receptor #3.  
XX  
KW Human; olfactory and pheromone G protein coupled; receptor; GPCR;  
KW tranquilizer; antidepressant; neuroleptic; endocrine; anabolic;  
KW anorectic; taste; fragrance; food additive; cosmetic; cell migration;  
KW sterility; psychotic disorder; neurological disorder; anxiety;  
KW schizophrenia; manic depression; depression; axonal growth;  
KW menstrual cycle; appetite sexual motivation; sexual attraction;  
XX  
OS Homo sapiens.  
XX  
PN WO200224726-A2.  
XX  
PD 28-MAR-2002.  
XX  
PF 21-SEP-2001; 2001WO-BE000162.  
XX  
PR 22-SEP-2000; 2000EP-00870211.  
XX  
PA (CHEM-) CHEMCOM SA.  
PI Velthen A;  
XX  
PI WPI; 2002-330013/36.  
DR N-PSDB; ABK68403.  
XX  
PT Novel pheromone G-protein coupled receptor and receptor-derived agonists,  
PT antagonists or inhibitors useful in food or cosmetic products or in the  
PT treatment or prevention of neurological disorders such as anxiety and  
PT schizophrenia.  
XX  
PS Disclosure; Page 23-25; 833bp; English.  
XX  
XX The invention relates to olfactory and pheromone G-protein coupled  
CC receptor (GPCR) or a protein 95% identical to the GPCR, a specific active  
CC portion and its encoding polynucleotide. Also included are an agonist,  
CC antagonist or inhibitor of the GPCR or the polynucleotide, a vector  
CC comprising the polynucleotide, a cell transformed by the vector, a non-  
CC human mammal comprising a partial or total deletion of the polynucleotide  
CC encoding the receptor and screening (detection and possibly, recovering)  
CC of compounds which are known or not known to be agonist, antagonists or  
CC inhibitors of natural compounds to the GPCR. The receptor-derived  
CC agonists, antagonists, inhibitors or compounds are used as an  
CC improvement, elimination or substitution of an existing taste and/or a  
CC fragrance of (or in) the food and/or cosmetic products. They can also be  
CC used in the preparation of medicament in the treatment and/or prevention  
CC of a mammalian disorder, such as cell migration, sterility, psychotic and  
CC neurological disorders, including anxiety, schizophrenia, manic  
CC depression, depression, for promoting axonal growth, nerve cell

CC connection and nerve regeneration for modulating male and female  
 CC endocrine functions, hormone production and the menstrual cycle, for the  
 CC prevention or the treatment by stimulation of several mammalian  
 CC behaviour, such as stimulation or suppression of appetite, sexual  
 CC motivation, sexual attraction, aggression and for promoting or  
 CC suppressing chemical communication between organisms. The present  
 CC sequence is a human olfactory and pheromone GPCR protein sequence  
 XX  
 SQ Sequence 326 AA;

Query Match 70.0%; Score 35; DB 5; Length 326;  
 Best Local Similarity 77.8%; Pred. No. 4.2e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 FLMFELDIV 9  
 Db 168 FLMFLIPIV 176

## RESULT 32

ADC85959 standard; protein; 326 AA.

AC ADC85959;  
 DT 01-JAN-2004 (first entry)  
 DE Human GPCR protein SEQ ID NO:412.  
 XX human; GPCR; guanosine triphosphate-binding protein coupled receptor;  
 KW gene therapy.  
 XX Homo sapiens.

EP1270724-A2.

02-JAN-2003.

18-JUN-2002; 2002EP-00013517.

18-JUN-2001; 2001JP-00246789.

(NAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
 (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

Suwa M, Asai K, Akiyama Y, Aburatani H;

WPI; 2003-315783/31.

N-PSDB; ADC85958.

PT New polynucleotide, useful for preparing a composition for treating a  
 patient in need of increased or suppressed activity or expression of the  
 PT guanosine triphosphate-binding protein coupled receptor.

PS Claim 2; SEQ ID NO 412; 28pp; English.

CC The invention relates to a novel polynucleotide encoding a guanosine  
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of  
 CC the invention may have a use in gene therapy. The polynucleotide and  
 CC polypeptide are useful for preparing a composition for treating a patient  
 CC in need of increased or suppressed activity or expression of the  
 CC guanosine triphosphate-binding protein coupled receptor. The protein  
 CC sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.

XX Sequence 326 AA;

Query Match 70.0%; Score 35; DB 7; Length 326;

Best Local Similarity 77.8%; Pred. No. 4.2e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 FLMFELDIV 9  
 Db 168 FLMFLIPIV 176

## RESULT 33

ADC85993 standard; protein; 326 AA.

AC ADC85993;  
 DT 01-JAN-2004 (first entry)  
 DE Human GPCR protein SEQ ID NO:446.  
 XX human; GPCR; guanosine triphosphate-binding protein coupled receptor;  
 KW gene therapy.  
 XX Homo sapiens.

EP1270724-A2.

02-JAN-2003.

18-JUN-2002; 2002EP-00013517.

18-JUN-2001; 2001JP-00246789.

(NAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
 (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

Suwa M, Asai K, Akiyama Y, Aburatani H;

WPI; 2003-315783/31.

N-PSDB; ADC85992.

PT New polynucleotide, useful for preparing a composition for treating a  
 patient in need of increased or suppressed activity or expression of the  
 PT guanosine triphosphate-binding protein coupled receptor.

PS Claim 2; SEQ ID NO 446; 28pp; English.

CC The invention relates to a novel polynucleotide encoding a guanosine  
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of  
 CC the invention may have a use in gene therapy. The polynucleotide and  
 CC polypeptide are useful for preparing a composition for treating a patient  
 CC in need of increased or suppressed activity or expression of the  
 CC guanosine triphosphate-binding protein coupled receptor. The protein  
 CC sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.

XX Sequence 326 AA;

Query Match 70.0%; Score 35; DB 7; Length 326;

Best Local Similarity 77.8%; Pred. No. 4.2e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 FLMFELDIV 9  
 Db 168 FLMFLIPIV 176

## RESULT 34

ADG83292 standard; protein; 326 AA.

ADG83292;

11-MAR-2004 (first entry)

Human Olfactory and pheromone GPCR #3.

XX Olfactory; pheromone; GPCR; G protein-coupled receptor; receptor; human;  
 KW flavour; detergent; soap; shampoo; fragrance; appetite control;  
 KW odour trap; receptor-mediated disorder; cell migration; cell death;  
 KW cell growth; psychotic disorder; neurological disorder; anxiety;  
 KW schizophrenia; manic depression; depression; wound modification.



XX OS Homo sapiens.  
 XX XX US2003221205-A1.  
 XX PN 27-NOV-2003.  
 XX PD 13-MAR-2003; 2003US-00387629.  
 XX PF 21-SEP-2001; 2001WO-BE000162.  
 XX PR (VEIT/) VEITHEN A.  
 XX PA Veithen A;  
 XX PI WPI; 2004-010906/01.  
 XX DR  
 XX PT New isolated pheromone G-protein coupled receptor, or its active portion,  
 PT useful for treating disorders affecting cell migration, death and growth,  
 PT or psychotic and neurological disorders.  
 XX PS  
 XX PS Disclosure; SEQ ID NO 6; 7pp; English.  
 XX CC The invention relates to an isolated olfactory and pheromone G protein-  
 CC coupled receptor (GPCR), or its active portion. Also included are an  
 CC isolated polynucleotide encoding the GPCR, a vector comprising the  
 CC polynucleotide, a cell comprising the vector, a non-human mammal  
 CC comprising a deletion of the polynucleotide, a transgenic non-human  
 CC mammal that overexpresses the polynucleotide, screening for compounds  
 CC that bind to the receptor polypeptide molecules, identifying a compound  
 CC as a ligand, agonist or antagonist of GPCR and a pharmaceutical  
 CC composition comprising the receptor, vector or cell, and a carrier. The  
 CC protein, polynucleotide, vector, cell and composition are useful for  
 CC preventing or treating receptor-mediated disorders, e.g. disorders  
 CC affecting cell migration, cell death, cell growth, or psychotic and  
 CC neurological disorders, including anxiety, schizophrenia, maniac  
 CC depression, depression or mood modification. They may also be useful in  
 CC developing detergents, soaps, shampoo, fragrances, in appetite control  
 CC and as an odour trap. The present sequence is an olfactory and pheromone  
 CC G protein-coupled receptor. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docID=20030221205.  
 CC XX  
 CC SQ Sequence 326 AA;  
 CC  
 CC Query Match 70.0%; Score 35; DB 8; Length 326;  
 CC Best Local Similarity 77.8%; Pred. No. 4.2e+02;  
 CC Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
 Db 168 FLWFLPIIV 176  
 |||||  
 |||||

RESULT 35  
 ADM42835  
 ID ADM42835 standard; protein; 379 AA.  
 XX  
 XX AC ADM42835;  
 XX  
 XX DT 03-JUN-2004 (first entry)  
 XX  
 XX DB Rat odourant receptor MOR106-2 SegID 79.  
 XX  
 XX KM rat; odourant receptor; chemical sensor system; taste; smell;  
 KM artificial sensory organ; olfactory stimulation; food industry;  
 KM hygiene inspection; environmental examination; disease diagnosis;  
 KM carvone.  
 XX  
 XX OS Rattus norvegicus.  
 XX  
 XX PN WO200310057-A1.

XX PD 04-DEC-2003.  
 XX XX 28-MAY-2003; 2003WO-JP006719.  
 XX PF 28-MAY-2002; 2002JP-00154239.  
 XX PR 13-JUN-2002; 2002JP-00172412.  
 XX PR 14-JAN-2003; 2003JP-00005175.  
 XX XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
 XX PA Sato T, Hirono J, Hamana H, Miyake M, Yoshikawa T, Miyake J;  
 XX PI WPI; 2004-023356/02.  
 XX DR N-PSDB; ADM42834.  
 XX XX  
 XX PT Chemical sensor systems based on chemical receptors introduced into cells  
 PT for immobilization onto support to form chip as component of sensor,  
 PT useful in detecting stimuli e.g. taste and smell applicable in food  
 PT industry.  
 XX PS  
 XX PS Disclosure; SEQ ID NO 79; 521pp; Japanese.  
 XX CC This invention relates to a novel chemical sensor system method.  
 CC Specifically, it refers to an isolated nucleic acid molecule that encodes  
 CC a receptor protein, which binds to chemicals that can stimulate the sense  
 CC of taste or smell for example. The present invention describes the  
 CC manufacture of a chip that acts as a support to immobilise transfected  
 CC cells expressing the receptor gene, such that this chip can be employed  
 CC as a component of the chemical sensor model. Furthermore, this chip is  
 CC useable as an artificial sensory organ where the chemical receptor  
 CC contains an olfactory receptor the sensor can react to olfactory  
 CC stimulation. Accordingly, these sensors are useful in the food industry  
 CC for analysing freshness of meat, fruit and vegetables, hygiene  
 CC inspection, environmental examination and disease diagnosis. Furthermore,  
 CC such systems are automatable for high throughput applications under  
 CC various conditions, even for differentiating optical isomers of R(-)-  
 CC carvone from S(+)-carvone easily. This polypeptide sequence is a rat  
 CC odourant receptor protein of the invention.  
 CC XX  
 CC SQ Sequence 379 AA;  
 CC  
 CC Query Match 70.0%; Score 35; DB 8; Length 379;  
 CC Best Local Similarity 55.6%; Pred. No. 5e+02;  
 CC Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9  
 Db 154 FLWFPVPII 162  
 |||||  
 |||||

RESULT 36  
 ADL04719  
 ID ADL04719 standard; protein; 543 AA.  
 XX  
 XX AC ADL04719;  
 XX  
 XX DT 06-MAY-2004 (first entry)  
 XX  
 XX DE M. catarrhalis protein #485.  
 XX  
 XX KM Moraxella catarrhalis; infection.  
 XX  
 XX OS Moraxella catarrhalis.  
 XX  
 XX PN US6673910-B1.  
 XX  
 XX PD 06-JAN-2004.  
 XX  
 XX PF 04-APR-2000; 2000US-00540236.  
 XX  
 XX PR 08-APR-1999; 99US-0128416P.  
 XX  
 XX PN

PA (GENO-) GENOME THERAPEUTICS CORP.  
 PI Breton GL;  
 XX  
 XX WPI; 2004-178127/17.  
 DR N-PSDB; ADL02799.  
 XX  
 PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for  
 PT preparing a composition for diagnosing, preventing or treating infection  
 PT caused by Moraxella catarrhalis.  
 XX  
 XX Disclousure; SEQ ID NO 2405; 423pp; English.  
 PS  
 CC The invention relates to an isolated nucleic acid encoding an Moraxella  
 CC catarrhalis polypeptide. The nucleic acid is useful for preparing a  
 CC composition for diagnosing, preventing or treating infection caused by  
 CC Moraxella catarrhalis. The present sequence represents the amino acid  
 CC sequence of a M. catarrhalis protein.  
 XX  
 SQ Sequence 543 AA;  
 Query Match 70.0%; Score 35; DB 8; Length 543;  
 Best Local Similarity 55.6%; Pred. No. 7.3e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 FLWFEIDIV 9  
 DB 308 FFWTEIDLI 316  
 RESULT 37  
 AAB25149  
 ID AAB25149 standard; protein; 599 AA.  
 XX  
 AC AAB25149;  
 DT 27-NOV-2000 (first entry)  
 XX  
 DE Eucalyptus grandis cell signalling involved protein SEQ ID NO:117.  
 XX  
 KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
 KW plant cell signalling; modulation; transgenic plant; pathogen; growth;  
 KW environmental change; development; cell proliferation; differentiation;  
 KW elongation; survival; disease resistance; nutrient metabolism.  
 XX  
 OS Eucalyptus grandis.  
 XX  
 PN WO200042171-A1.  
 XX  
 PD 20-JUL-2000.  
 XX  
 PF 11-JAN-2000; 2000MO-US000724.  
 XX  
 PR 12-JAN-1999; 99US-00228986.  
 PR 01-NOV-1999; 99US-0162866P.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PI Strabala TJ, Nieuwenhuizen NJ;  
 XX  
 DR WPI; 2000-476052/41.  
 XX  
 PT Isolated polynucleotide encoding a polypeptide involved in cell signalling  
 PT used for generating transgenic plants with modified responses to external  
 PT signals.  
 XX  
 PS Claim 3; Page 98-100; 527pp; English.  
 XX  
 CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide  
 CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
 CC pine (Pinus radiata also known as Monterey pine). The protein sequences  
 CC are involved in cell signalling. The polynucleotide and protein sequences  
 CC can be used to modify the response of plant cells to external signals

CC e.g. environmental changes or pathogens during the growth and development  
 CC of a plant. They can be used to modify cell proliferation,  
 CC differentiation, elongation and survival, resistance to disease and  
 CC nutrient metabolism. Examples of modifications which can be produced are  
 CC altered fruit ripening and senescence of leaves and flowers e.g. to delay  
 CC senescence and prolong the life of cut flowers or enhance senescence of  
 CC reproductive organs to engineer sterile plants. Other modifications can  
 CC be used to delay senescence in selected cell types or organs providing  
 CC fruit and vegetables which have a longer shelf life between harvest and  
 CC consumption, or to decrease branching frequency in forest tree species  
 CC giving long stretches of valuable knot-free clear wood which can be used  
 CC in solid timber furniture and veneers  
 XX  
 SQ Sequence 599 AA;  
 Query Match 70.0%; Score 35; DB 3; Length 599;  
 Best Local Similarity 83.3%; Pred. No. 8.1e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LWFED 7  
 DB 33 LWFED 38  
 RESULT 38  
 AAB25508  
 ID AAB25508 standard; protein; 816 AA.  
 XX  
 AC AAB25508;  
 DT 27-NOV-2000 (first entry)  
 XX  
 DE Eucalyptus grandis cell signalling involved protein SEQ ID NO:827.  
 XX  
 KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
 KW plant cell signalling; modulation; transgenic plant; pathogen; growth;  
 KW environmental change; development; cell proliferation; differentiation;  
 KW elongation; survival; disease resistance; nutrient metabolism.  
 XX  
 OS Eucalyptus grandis.  
 XX  
 PN WO200042171-A1.  
 XX  
 PD 20-JUL-2000.  
 XX  
 PF 11-JAN-2000; 2000MO-US000724.  
 XX  
 PR 12-JAN-1999; 99US-00228986.  
 PR 01-NOV-1999; 99US-0162866P.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PI Strabala TJ, Nieuwenhuizen NJ;  
 XX  
 DR WPI; 2000-476052/41.  
 XX  
 PT Isolated polynucleotide encoding a polypeptide involved in cell signalling  
 PT used for generating transgenic plants with modified responses to external  
 PT signals.  
 XX  
 PS Claim 3; Page 394-395; 527pp; English.  
 XX  
 CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide  
 CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
 CC pine (Pinus radiata also known as Monterey pine). The protein sequences  
 CC are involved in cell signalling. The polynucleotide and protein sequences  
 CC can be used to modify the response of plant cells to external signals  
 CC e.g. environmental changes or pathogens during the growth and development  
 CC of a plant. They can be used to modify cell proliferation,  
 CC differentiation, elongation and survival, resistance to disease and  
 CC nutrient metabolism. Examples of modifications which can be produced are  
 CC altered fruit ripening and senescence of leaves and flowers e.g. to delay  
 CC senescence and prolong the life of cut flowers or enhance senescence of

CC reproductive organs to engineer sterile plants. Other modifications can  
CC be used to delay senescence in selected cell types or organs providing  
CC fruit and vegetables which have a longer shelf life between harvest and  
CC consumption, or to decrease branching frequency in forest tree species  
CC giving long stretches of valuable knot-free clear wood which can be used  
CC in solid timber furniture and veneers  
XX

SO Sequence 816 AA;

Query Match 70.0%; Score 35; DB 3; Length 816;  
Best Local Similarity 83.3%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LWFELD 7  
DB 250 LWFELD 255

RESULT 39

ADA33169 ADA33169 standard; protein; 951 AA.

AC ADA33169;

DT 20-NOV-2003 (first entry)

DE Acinetobacter baumannii protein #330.

KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;  
KW plant biocontrol agent.

OS Acinetobacter baumannii.

PN US6562958-B1.

PD 13-MAY-2003.

PP 04-JUN-1999; 99US-00328352.

PR 09-JUN-1998; 98US-0088701P.

PA (GENO-) GENOME THERAPEUTICS CORP.

P1 Breton G, Bush D;

DR WPI; 2003-576092/54.

DR N-PSDB; ADA29043.

PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
PT for diagnosing a bacterial disease, as components of antibacterial  
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
PT plants.

PS Example; SEQ ID NO 4456; 328pp; English.

CC The invention relates to isolated Acinetobacter baumannii nucleic acids.  
CC The A. baumannii nucleic acids and polypeptides are useful as reagents  
CC for diagnosing a bacterial disease, as components of antibacterial  
CC vaccines, as targets for antibacterial drugs, to detect the presence of  
CC A. baumannii and other Acinetobacter species in a sample, in screening  
CC compounds for the ability to interfere with the A. baumannii life cycle  
CC or to inhibit A. baumannii infection, and as biocontrol agents for  
CC plants. The present sequence represents the amino acid sequence of an A.  
CC baumannii protein.  
XX

SO Sequence 951 AA;

Query Match 70.0%; Score 35; DB 6; Length 951;

Best Local Similarity 50.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 LWFELD 9  
DB 250 LWFELD 255

DB 801 WFKVDVV 808

RESULT 40

AAM88375 AAM88375 standard; protein; 37 AA.

AC AAM88375;

DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen SEQ ID NO:15968.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis.

OS Homo sapiens.

PN WO200157182-A2.

PD 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US001354.

PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214866P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 23-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-SEP-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

[illegible]

PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Barash SC, Ruben SM;	
PI	WPI; 2001-483426/52.	
DR	N-PSDB; AAK61156.	
DX		
XX		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PR	useful for preventing, diagnosing and/or treating cancers and metastasis.	
XX		
PS	Claim 11; SEQ ID NO 15968; 3071pp + Sequence Listing; English.	
XX		
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
CC	amino acid sequences given in AAM62170 to AAM91921. (I) have cytostatic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I). By expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (II)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/haematopoietic-related diseases, especially	
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703	
CC	to AAK67694 represent human immune/haematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM62169	
CC	represent sequences used in the exemplification of the present invention	
XX		
SO	Sequence 37 AA:	
QY	Query Match	68.0%; Score 34; DB 4; Length 37;
	Best Local Similarity	33.3%; Pred. No. 62;
	Matches	3; Conservative
		6; Mismatches
		0; Indels
		0; Gaps
DB	17 FLNMOLULEV1	25

Search completed: January 12, 2005, 20:06:53  
Job time : 205.6 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2005, 20:16:00 ; Search time 93.6 Seconds  
(without alignments)  
34.739 Million cell updates/sec

Title: US-09-870-216C-7  
Perfect score: 50  
Sequence: 1 FLWPEIDIV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PTCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PTCT\_NEW\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	9	11 US-09-870-216C-7	Sequence 7, Appl1
2	50	100.0	9	13 US-10-017-327-7	Sequence 7, Appl1
3	39	78.0	94	17 US-10-425-115-206553	Sequence 206553,
4	39	78.0	1717	14 US-10-126-339-2	Sequence 2, Appl1
5	39	78.0	1717	14 US-10-126-148-2	Sequence 2, Appl1
6	39	78.0	1717	17 US-10-798-058-8	Sequence 8, Appl1
7	38	76.0	112	15 US-10-424-599-183462	Sequence 183462,
8	37	74.0	529	14 US-10-369-493-18446	Sequence 18446, A
9	36	72.0	58	15 US-10-424-599-144215	Sequence 144215,
10	36	72.0	402	14 US-10-369-493-20052	Sequence 20052, A
11	36	72.0	700	16 US-10-408-765A-2042	Sequence 2042, Ap
12	36	72.0	1759	17 US-10-798-058-2	Sequence 2, Appl1
13	35	70.0	37	15 US-10-424-599-273756	Sequence 273756,

14	35	70.0	210	15	US-10-424-599-236699	Sequence 236699,
15	35	70.0	217	15	US-10-041-615-63	Sequence 63, Appl1
16	35	70.0	217	15	US-10-072-012-557	Sequence 557, App
17	35	70.0	232	14	US-10-017-161-1002	Sequence 1002, Ap
18	35	70.0	232	14	US-10-292-798-852	Sequence 852, App
19	35	70.0	315	15	US-10-041-615-16	Sequence 16, Appl1
20	35	70.0	315	15	US-10-072-012-192	Sequence 192, App
21	35	70.0	326	14	US-10-017-161-460	Sequence 460, App
22	35	70.0	326	14	US-10-017-161-506	Sequence 506, App
23	35	70.0	326	14	US-10-387-629-6	Sequence 6, Appl1
24	35	70.0	326	14	US-10-292-798-412	Sequence 412, App
25	35	70.0	326	14	US-10-292-798-446	Sequence 446, App
26	35	70.0	326	15	US-10-343-650A-640	Sequence 640, App
27	35	70.0	599	14	US-10-101-664A-117	Sequence 117, App
28	35	70.0	720	15	US-10-425-114-72890	Sequence 72890, A
29	35	70.0	816	14	US-10-101-664A-827	Sequence 827, App
30	35	70.0	1240	14	US-10-101-664A-976	Sequence 976, App
31	34	68.0	44	17	US-10-425-115-227895	Sequence 227895,
32	34	68.0	51	9	US-09-798-029-21	Sequence 21, Appl1
33	34	68.0	61	16	US-10-437-963-147599	Sequence 147599,
34	34	68.0	62	17	US-10-425-115-363307	Sequence 363307,
35	34	68.0	75	15	US-10-424-599-218678	Sequence 218678,
36	34	68.0	77	15	US-10-424-599-152779	Sequence 152779,
37	34	68.0	175	15	US-10-424-599-264160	Sequence 264160,
38	34	68.0	175	15	US-10-424-599-264161	Sequence 264161,
39	34	68.0	175	15	US-10-424-599-261009	Sequence 261009,
40	34	68.0	218	15	US-10-424-599-264158	Sequence 264158,
41	34	68.0	296	15	US-10-282-122A-53995	Sequence 53995, A
42	34	68.0	378	16	US-10-159-257A-176	Sequence 176, App
43	34	68.0	393	15	US-10-424-599-177498	Sequence 177498,
44	34	68.0	416	15	US-10-282-122A-48716	Sequence 48716, A
45	34	68.0	422	15	US-10-471-243-6	Sequence 6, Appl1

## ALIGNMENTS

RESULT 1  
US-09-870-216C-7  
; Sequence 7, Application US/09870216C  
; Publication No. US20040138135A1  
; GENERAL INFORMATION:  
; APPLICANT: Charles A. Nicolette  
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
; FILE REFERENCE: 68126881210100  
; CURRENT APPLICATION NUMBER: US/09/870,216C  
; CURRENT FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: 60/209,391  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/226,256  
; PRIOR FILING DATE: 2000-08-17  
; PRIOR APPLICATION NUMBER: 60/257,008  
; PRIOR FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-870-216C-7

Query Match  
Best Local Similarity 100.0%; Score 50; DB 11; Length 9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 FLWPEIDIV 9  
RESULT 2  
US-10-017-327-7  
; Sequence 7, Application US/10017327

Publication No. US20020155471A1  
GENERAL INFORMATION:  
APPLICANT: Charles A. Nicolette  
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND  
METHODS FOR USING SAME  
FILE REFERENCE: GZ 2101.20  
CURRENT APPLICATION NUMBER: US/10/017,337  
CURRENT FILING DATE: 2001-12-06  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-017-327-7

Query Match 100.0%; Score 50; DB 13; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMFEIDIV 9  
Db 1 FLMFEIDIV 9

RESULT 3  
US-10-425-115-206553  
Sequence 206553, Application US/10425115  
Publication No. US20040214272A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants  
FILE REFERENCE: 38-21(53222)B  
CURRENT APPLICATION NUMBER: US/10/425,115  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 369326  
SEQ ID NO 206553  
LENGTH: 94  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(94)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_119961C.1.pap  
US-10-425-115-206553

Query Match 78.0%; Score 39; DB 17; Length 94;  
Best Local Similarity 55.6%; Pred. No. 27;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEIDIV 9  
Db 26 FLMFEIDIV 34

RESULT 4  
US-10-126-339-2  
Sequence 2, Application US/10126339  
Publication No. US20030166891A1  
GENERAL INFORMATION:  
APPLICANT: Flanagan, Ronald D.  
APPLICANT: Mathis, John P.  
APPLICANT: Meyer, Terry E.  
TITLE OF INVENTION: No. US20030166891A1 Bt Toxin Receptors From  
TITLE OF INVENTION: Lepidopteran Insects and Methods of Use  
FILE REFERENCE: 35718/204664  
CURRENT APPLICATION NUMBER: US/10/126,339

CURRENT FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: US/09/715,909  
PRIOR FILING DATE: 2000-11-17  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1717  
TYPE: PRT  
ORGANISM: Ostrinia nubilalis  
US-10-126-339-2

Query Match 78.0%; Score 39; DB 14; Length 1717;  
Best Local Similarity 85.7%; Pred. No. 4.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMFEID 7  
Db 650 YLMFEID 656

RESULT 5  
US-10-126-148-2  
Sequence 2, Application US/10126148  
Publication No. US20030170787A1  
GENERAL INFORMATION:  
APPLICANT: Flanagan, Ronald D.  
APPLICANT: Mathis, John P.  
APPLICANT: Meyer, Terry E.  
TITLE OF INVENTION: No. US20030170787A1 Bt Toxin Receptors From  
TITLE OF INVENTION: Lepidopteran Insects and Methods of Use  
FILE REFERENCE: 35718/204664  
CURRENT APPLICATION NUMBER: US/10/126,148  
CURRENT FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: US/09/715,909  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/166,285  
PRIOR FILING DATE: 1999-11-18  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1717  
TYPE: PRT  
ORGANISM: Ostrinia nubilalis  
US-10-126-148-2

Query Match 78.0%; Score 39; DB 14; Length 1717;  
Best Local Similarity 85.7%; Pred. No. 4.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMFEID 7  
Db 650 YLMFEID 656

RESULT 6  
US-10-798-058-8  
Sequence 8, Application US/10798058  
Publication No. US20040209329A1  
GENERAL INFORMATION:  
APPLICANT: Mathis, John P.  
TITLE OF INVENTION: Novel Bt Toxin Receptors and Methods of  
TITLE OF INVENTION: Use  
FILE REFERENCE: 35718/274644  
CURRENT APPLICATION NUMBER: US/10/798,058  
CURRENT FILING DATE: 2004-03-11  
PRIOR APPLICATION NUMBER: US 60/455,085  
PRIOR FILING DATE: 2003-03-14  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 1717  
TYPE: PRT  
ORGANISM: Ostrinia nubilalis

US-10-798-058-8

Query Match 78.0%; Score 39; DB 17; Length 1717;  
Best Local Similarity 85.7%; Pred. No. 4.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLWFEID 7  
Db 650 YLWFEID 656

RESULT 7

US-10-424-599-183462  
; Sequence 183462, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 183462  
; LENGTH: 112  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(112)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_13667C.1.pep  
US-10-424-599-183462

Query Match 76.0%; Score 38; DB 15; Length 112;  
Best Local Similarity 77.8%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
Db 39 FLWFEIDIV 47

RESULT 8

US-10-369-493-18446  
; Sequence 18446, Application US/10369493  
; Publication No. US2003023675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 18446  
; LENGTH: 529  
; TYPE: PRT  
; ORGANISM: Lactococcus lactis  
US-10-369-493-18446

Query Match 74.0%; Score 37; DB 14; Length 529;  
Best Local Similarity 66.7%; Pred. No. 3.1e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
Db 49 FLWFEIDIV 57

RESULT 9

US-10-424-599-144215  
; Sequence 144215, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 144215  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_101239C.1.pep  
US-10-424-599-144215

Query Match 72.0%; Score 36; DB 15; Length 58;  
Best Local Similarity 75.0%; Pred. No. 58;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FLWFEIDIV 9  
Db 45 FLWFEIDIV 52

RESULT 10

US-10-369-493-20052  
; Sequence 20052, Application US/10369493  
; Publication No. US2003023675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 20052  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: No. US2003023675A1loc punctiforme  
US-10-369-493-20052

Query Match 72.0%; Score 36; DB 14; Length 402;  
Best Local Similarity 71.4%; Pred. No. 3.6e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFEIDIV 9  
Db 245 WFEIDIV 251

RESULT 11

```
US-10-408-765A-2042
; Sequence 2042, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Robin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2042
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2042

Query Match          72.0%; Score 36; DB 16; Length 700;
Best Local Similarity 75.0%; Pred. No. 6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEID 8
DB 612 FLMFENDM 619

RESULT 12
US-10-798-058-2
; Sequence 2, Application US/10798058
; Publication No. US20040209329A1
; GENERAL INFORMATION:
; APPLICANT: Mathis, John P.
; TITLE OF INVENTION: Novel Bt Toxin Receptors and Methods of
; FILE REFERENCE: 35718/274644
; CURRENT APPLICATION NUMBER: US/10/798,058
; CURRENT FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: US 60/455,085
; PRIOR FILING DATE: 2003-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1759
; TYPE: PRT
; ORGANISM: Agrotis ipsilon
US-10-798-058-2

Query Match          72.0%; Score 36; DB 17; Length 1759;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LWFED 7
DB 668 LWFED 673

RESULT 13
US-10-424-599-273756
; Sequence 273756, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
```

```
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 265684
; SEQ ID NO 273756
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89222C.1.pcp
US-10-424-599-273756

Query Match          70.0%; Score 35; DB 15; Length 37;
Best Local Similarity 71.4%; Pred. No. 57;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLMFEID 7
DB 13 FLMFENDM 19

RESULT 14
US-10-424-599-236699
; Sequence 236699, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 236699
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(210)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_55766C.1.pcp
US-10-424-599-236699

Query Match          70.0%; Score 35; DB 15; Length 210;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEID 7
DB 55 FLMFENDM 61

RESULT 15
US-10-041-615-63
; Sequence 63, Application US/10041615
; Publication No. US20040014038A1
; GENERAL INFORMATION:
; APPLICANT: Casman, Stracie J
; APPLICANT: Edinger, Shlomil R
; APPLICANT: Ellerman, Karen
; APPLICANT: Smithson, Glenda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Padigar, Muralidhara
; TITLE OF INVENTION: No. US20040014038A1 GPCR-Like Proteins and Nucleic Acids Encodi
; FILE REFERENCE: 21402-233-061
; CURRENT APPLICATION NUMBER: US/10/041,615
; CURRENT FILING DATE: 2003-01-29
```



```

; PRIOR APPLICATION NUMBER: 60/259,552
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/260,544
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/277,405
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 63
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-615-63

Query Match          70.0%; Score 35; DB 15; Length 217;
Best Local Similarity 77.8%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLWFEIDIV 9
DB      86 FLWFLIPIV 94

RESULT 16
US-10-072-012-557
; Sequence 557, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patursajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Raetelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Albrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
```

```

; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 557
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-557

Query Match          70.0%; Score 35; DB 15; Length 217;
Best Local Similarity 77.8%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLWFEIDIV 9
DB      86 FLWFLIPIV 94

RESULT 17
US-10-017-161-1002
; Sequence 1002, Application US/10017161
; Publication No. US2003014368A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIRO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABRATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 08435/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1002
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-1002

Query Match          70.0%; Score 35; DB 14; Length 232;
Best Local Similarity 77.8%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLWFEIDIV 9
DB      168 FLWFLIPIV 176

RESULT 18
US-10-292-798-852
; Sequence 852, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIRO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABRATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 08435/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 852
; LENGTH: 232
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-852

Query Match      70.0%; Score 35; DB 14; Length 232;
Best Local Similarity 77.8%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLMFELDIV 9
        |||||
Db      168 FLMFLIPIV 176

RESULT 19
US-10-041-615-16
; Sequence 16, Application US/10041615
; Publication No. US20040014038A1
; GENERAL INFORMATION:
; APPLICANT: Casman, Stacie J
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Smithson, Glenda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: No. US20040014038A1el GPCR-Like Proteins and Nucleic Acids Encodi
; FILE REFERENCE: 21402-233-061
; CURRENT APPLICATION NUMBER: US/10/041,615
; PRIOR FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/259,552
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/260,544
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/277,405
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: CuroSeqdist version 0.1
; SEQ ID NO 16
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-615-16

Query Match      70.0%; Score 35; DB 15; Length 315;
Best Local Similarity 77.8%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLMFELDIV 9
        |||||
Db      157 FLMFLIPIV 165

RESULT 20
US-10-072-012-192
; Sequence 192, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchervet, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimr Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
```

```

; APPLICANT: Pirtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 192
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-192

Query Match      70.0%; Score 35; DB 15; Length 315;
Best Local Similarity 77.8%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 FLMFELDIV 9
        |||||
Db      157 FLMFLIPIV 165

RESULT 21
US-10-017-161-460
; Sequence 460, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 460
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-460

Query Match      70.0%; Score 35; DB 14; Length 326;
Best Local Similarity 77.8%; Pred. No. 4.4e+02;
```

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
||| |  
Db 168 FLWFLIPDIV 176

RESULT 22  
US-10-017-161-506  
; Sequence 506, Application US/10017161  
; Publication No. US20030143668A1  
; GENERAL INFORMATION:  
; APPLICANT: SUMA, MAKIKO  
; APPLICANT: ASAI, KIYOSHI  
; APPLICANT: ABIYAMA, YUTAKA  
; APPLICANT: ABIYAMA, YUTAKA  
; APPLICANT: ABIYAMA, YUTAKA  
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS  
; FILE REFERENCE: 084335/0152  
; CURRENT FILING DATE: 2002-12-18  
; PRIOR APPLICATION NUMBER: JP 2001-246789  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 2430  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 506  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-017-161-506

Query Match  
Best Local Similarity 77.0%; Score 35; DB 14; Length 326;  
Best Local Similarity 77.8%; Pred. No. 4.4e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
||| |  
Db 168 FLWFLIPDIV 176

RESULT 23  
US-10-387-629-6  
; Sequence 6, Application US/10387629  
; Publication No. US20030221205A1  
; GENERAL INFORMATION:  
; APPLICANT: Veithen, Alex  
; APPLICANT: Chemcom S.A.  
; TITLE OF INVENTION: Olfactory and Pheromones G-Protein coupled Receptors  
; FILE REFERENCE: 9409/2192  
; CURRENT APPLICATION NUMBER: US/10/387,629  
; CURRENT FILING DATE: 2003-03-13  
; NUMBER OF SEQ ID NOS: 254  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-387-629-6

Query Match  
Best Local Similarity 77.0%; Score 35; DB 14; Length 326;  
Best Local Similarity 77.8%; Pred. No. 4.4e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
||| |  
Db 168 FLWFLIPDIV 176

RESULT 24  
US-10-292-798-412  
; Sequence 412, Application US/10292798  
; Publication No. US20030235833A1  
; GENERAL INFORMATION:  
; APPLICANT: SUMA, MAKIKO

; APPLICANT: ASAI, KIYOSHI  
; APPLICANT: ABIYAMA, YUTAKA  
; APPLICANT: ABIYAMA, YUTAKA  
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS  
; FILE REFERENCE: 084335/166  
; CURRENT APPLICATION NUMBER: US/10/292,798  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: 10/017,161  
; PRIOR FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: JP 2001-246789  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 2070  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 412  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-292-798-412

Query Match  
Best Local Similarity 77.0%; Score 35; DB 14; Length 326;  
Best Local Similarity 77.8%; Pred. No. 4.4e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
||| |  
Db 168 FLWFLIPDIV 176

RESULT 25  
US-10-292-798-446  
; Sequence 446, Application US/10292798  
; Publication No. US20030235833A1  
; GENERAL INFORMATION:  
; APPLICANT: SUMA, MAKIKO  
; APPLICANT: ASAI, KIYOSHI  
; APPLICANT: ABIYAMA, YUTAKA  
; APPLICANT: ABIYAMA, YUTAKA  
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS  
; FILE REFERENCE: 084335/166  
; CURRENT APPLICATION NUMBER: US/10/292,798  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: 10/017,161  
; PRIOR FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: JP 2001-246789  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 2070  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 446  
; LENGTH: 326  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-292-798-446

Query Match  
Best Local Similarity 77.0%; Score 35; DB 14; Length 326;  
Best Local Similarity 77.8%; Pred. No. 4.4e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLWFEIDIV 9  
||| |  
Db 168 FLWFLIPDIV 176

RESULT 26  
US-10-343-650A-640  
; Sequence 640, Application US/10343650A  
; Publication No. US20040067499A1  
; GENERAL INFORMATION:  
; APPLICANT: HAGA, TATSUYA  
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR  
; FILE REFERENCE: 31671-186347  
; CURRENT APPLICATION NUMBER: US/10/343,650A  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: JP 2000/237818

;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: JP 2001/34434  
;; PRIOR FILING DATE: 2001-02-13  
;; NUMBER OF SEQ ID NOS: 694  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO: 640  
;; LENGTH: 326  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-343-650A-640

Query Match 70.0%; Score 35; DB 15; Length 326;  
Best Local Similarity 77.8%; Pred. No. 4,4e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LWFEPIDV 9  
Db 168 LWFEPIDV 176

RESULT 27  
US-10-101-464A-117  
; Sequence 117, Application US/10101464A  
; Publication No. US20030046728A1  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; FILE REFERENCE: 11000.1020c2  
; CURRENT APPLICATION NUMBER: US/10/101,464A  
; PRIOR FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT/US00/00724  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 117  
; LENGTH: 599  
; TYPE: PRT  
; ORGANISM: Eucalyptus grandis  
US-10-101-464A-117

Query Match 70.0%; Score 35; DB 14; Length 599;  
Best Local Similarity 83.3%; Pred. No. 7,8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFEPID 7  
Db 33 LWFEPID 38

RESULT 28  
US-10-425-114-72890  
; Sequence 72890, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114

;; CURRENT FILING DATE: 2003-04-28  
;; NUMBER OF SEQ ID NOS: 73128  
;; SEQ ID NO: 72890  
;; LENGTH: 720  
;; TYPE: PRT  
;; ORGANISM: Arabidopsis thaliana columbia  
;; FEATURE:  
;; OTHER INFORMATION: Clone ID: JC-ATXLIB327417P1D12\_F11.pep  
US-10-425-114-72890

Query Match 70.0%; Score 35; DB 15; Length 720;  
Best Local Similarity 57.1%; Pred. No. 9,3e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LWFEPID 7  
Db 528 YWFEPVD 534

RESULT 29  
US-10-101-464A-827  
; Sequence 827, Application US/10101464A  
; Publication No. US20030046728A1  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; FILE REFERENCE: 11000.1020c2  
; CURRENT APPLICATION NUMBER: US/10/101,464A  
; PRIOR FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT/US00/00724  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 827  
; LENGTH: 816  
; TYPE: PRT  
; ORGANISM: Eucalyptus grandis  
US-10-101-464A-827

Query Match 70.0%; Score 35; DB 14; Length 816;  
Best Local Similarity 83.3%; Pred. No. 1e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFEPID 7  
Db 250 LWFEPVD 255

RESULT 30  
US-10-101-464A-976  
; Sequence 976, Application US/10101464A  
; Publication No. US20030046728A1  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; FILE REFERENCE: 11000.1020c2  
; CURRENT APPLICATION NUMBER: US/10/101,464A  
; PRIOR FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986

;; PRIOR FILING DATE: 1999-01-12  
;; PRIOR APPLICATION NUMBER: 60/162,866  
;; PRIOR FILING DATE: 1999-11-01  
;; PRIOR APPLICATION NUMBER: PCT/US00/00724  
;; PRIOR FILING DATE: 2000-01-11  
;; NUMBER OF SEQ ID NOS: 989  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 976  
;; LENGTH: 1240  
;; TYPE: PRT  
;; ORGANISM: Eucalyptus grandis  
US-10-101-464A-976

Query Match 70.0%; Score 35; DB 14; Length 1240;  
Best Local Similarity 83.3%; Pred. No. 1.5e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMFEID 7  
DB 674 LMFEVD 679

RESULT 31  
US-10-425-115-227895  
; Sequence 227895, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yinhua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 227895  
; LENGTH: 44  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_139435C.1.pdp  
US-10-425-115-227895

Query Match 68.0%; Score 34; DB 17; Length 44;  
Best Local Similarity 83.3%; Pred. No. 1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMFEI 6  
DB 30 LMFEEL 35

RESULT 32  
US-09-798-029-21  
; Sequence 21, Application US/09798029  
; Patent No. US20020019030A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel  
; APPLICANT: Macbeth, Kyle J.  
; APPLICANT: Rudolph-Owen, Laura  
; TITLE OF INVENTION: 25324, 50287, 28899, 47007, and 42967  
; FILE REFERENCE: 38155-20001.00  
; CURRENT APPLICATION NUMBER: US/09/798,029  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: US 60/185,711  
; PRIOR FILING DATE: 2001-02-27  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21

;; LENGTH: 51  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Consensus amino acid  
US-09-798-029-21

Query Match 68.0%; Score 34; DB 9; Length 51;  
Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMFEI 6  
DB 46 LMFEI 51

RESULT 33  
US-10-437-963-147599  
; Sequence 147599, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yinhua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 147599  
; LENGTH: 61  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_48112C.1.pdp  
US-10-437-963-147599

Query Match 68.0%; Score 34; DB 16; Length 61;  
Best Local Similarity 62.5%; Pred. No. 1.4e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LMFEIDI 8  
DB 34 FRWFSIDV 41

RESULT 34  
US-10-425-115-363307  
; Sequence 363307, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yinhua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 363307  
; LENGTH: 62  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure

```
LOCATION: (1)..(62)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_94508C.1.pep
US-10-425-115-363307

Query Match
Best Local Similarity 68.0%; Score 34; DB 15; Length 62;
Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLMFEIDIV 9
Db 35 FVWKEIDSV 43

RESULT 35
US-10-424-599-218678
Sequence 218678, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
Title of Invention: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 218678
LENGTH: 75
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(75)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_39495C.1.pep
US-10-424-599-218678

Query Match
Best Local Similarity 68.0%; Score 34; DB 15; Length 75;
Pred. No. 1.7e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEIDIV 9
Db 2 FLMDDIDVL 10

RESULT 36
US-10-424-599-152779
Sequence 152779, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
Title of Invention: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 152779
LENGTH: 77
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(77)
```

```
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_108984C.1.pep
US-10-424-599-152779

Query Match
Best Local Similarity 68.0%; Score 34; DB 15; Length 77;
Pred. No. 1.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 WFEID 8
Db 4 WFEVDV 9

RESULT 37
US-10-424-599-264160
Sequence 264160, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
Title of Invention: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 264160
LENGTH: 175
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_80558C.1.pep
US-10-424-599-264160

Query Match
Best Local Similarity 68.0%; Score 34; DB 15; Length 175;
Pred. No. 3.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEID 7
Db 106 FLMFNDV 112

RESULT 38
US-10-424-599-264161
Sequence 264161, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
Title of Invention: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 264161
LENGTH: 175
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_80558C.1.pep
US-10-424-599-264161

Query Match
Best Local Similarity 68.0%; Score 34; DB 15; Length 175;
Pred. No. 3.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

OY 1 FLWFEID 7  
|:|:|:  
Db 106 FLMXHVND 112

Search completed: January 12, 2005, 20:48:27  
Job time : 96 secs

RESULT 39  
US-10-424-599-261009

; Sequence 261009, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 261009  
; LENGTH: 178  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(178)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_77713C.1.pep  
US-10-424-599-261009

Query Match 68.0%; Score 34; DB 15; Length 178;  
Best Local Similarity 55.6%; Pred. No. 3.7e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 FLWFEIDIV 9  
|:|:|:  
Db 41 FLMXHVNDIL 49

RESULT 40  
US-10-424-599-264158  
; Sequence 264158, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 264158  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_80556C.1.pep  
US-10-424-599-264158

Query Match 68.0%; Score 34; DB 15; Length 215;  
Best Local Similarity 57.1%; Pred. No. 4.5e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLWFEID 7  
|:|:|:  
Db 106 FLMXHVND 112

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:46:14 ; Search time 16.4 Seconds  
(without alignments)  
52.802 Million cell updates/sec

Title: US-09-870-216C-7  
Perfect score: 50  
Sequence: 1 FLWFEIDIV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_79:.\*  
1: p1r1:.\*  
2: p1r2:.\*  
3: p1r3:.\*  
4: p1r4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	78.0	474	2 T38905	probable valine-py
2	39	78.0	587	2 A10180	conserved hypothet
3	37	74.0	110	2 S40136	T-cell receptor V-
4	37	74.0	134	2 C29774	T-cell receptor al
5	37	74.0	150	2 AG1585	hypothetical prote
6	37	74.0	150	2 AH1231	hypothetical prote
7	37	74.0	529	2 B86815	hypothetical prote
8	36	72.0	269	2 A86641	hypothetical prote
9	36	72.0	413	2 B64815	yhho protein - Bsc
10	36	72.0	413	2 D85587	probable synthetas
11	36	72.0	413	2 C90737	cardiolipin synthe
12	35	70.0	250	2 D69457	hypothetical prote
13	35	70.0	280	2 B82372	Spoom-related prot
14	35	70.0	648	2 T08856	hypothetical prote
15	35	70.0	848	2 T25325	hypothetical prote
16	35	70.0	1190	2 T00842	probable histidini
17	35	70.0	1207	2 T52459	sensory transducti
18	34	68.0	121	2 A12244	dihydronoepterin a
19	34	68.0	175	2 AB0185	probable membrane
20	34	68.0	209	2 F16162	glycosyl transfera
21	34	68.0	229	2 AB1458	ribosomal protein
22	34	68.0	254	2 C90024	hypothetical prote
23	34	68.0	422	2 T05592	lysine transamin
24	34	68.0	469	2 S44620	C50C3.1 protein -
25	34	68.0	643	2 S55593	membrane protein S
26	34	68.0	912	2 T49399	hypothetical prote
27	34	68.0	2363	2 T38841	probable pre-mRNA
28	33	66.0	55	2 C82565	hypothetical prote
29	33	66.0	126	2 A82519	hypothetical prote

30	33	66.0	143	2 A11511	hypothetical prote
31	33	66.0	150	1 S76177	dihydronoepterin a
32	33	66.0	190	2 A82343	conserved hypothet
33	33	66.0	218	2 G86612	hypothetical prote
34	33	66.0	226	2 A86316	protein T10022.3 l
35	33	66.0	269	2 H82616	hypothetical prote
36	33	66.0	269	2 D95396	probable polyketid
37	33	66.0	283	2 B11876	probable polypopr
38	33	66.0	284	1 C64639	polipoprotein dia
39	33	66.0	285	2 AG2429	delta-9 desaturase
40	33	66.0	343	2 F72011	hypothetical prote
41	33	66.0	461	2 G64206	hexosephosphate tr
42	33	66.0	470	2 T41409	probable valine-py
43	33	66.0	507	2 S75961	NADH2 dehydrogenas
44	33	66.0	516	2 S42093	cellulose 1,4-beta
45	33	66.0	546	2 A69484	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

T38905 Probable valine-pyruvate transaminase (EC 2.6.1.66) SPAC56E4.03 [similarity] - fission y

C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C/Accession: T38905

R/Conor: R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, September 1997

A/Accession: T38905

A/Status: translated from GB/EMBL/DBDJ

A/Molecule type: DNA

A/Residues: 1-474 <CON>

A/Cross-references: UNIPROT:O14192; EMBL:Z99261; NID:G4038621; PTDN:CAH16394.1; PTD:G241

A/Experimental source: Strain 972h-; cosmid C56E4

C/Genetics:

A/Map position: 1

A/Map position: 1

C/Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate

F1297/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 78.0%; Score 39; DB 2; Length 474;  
Best Local Similarity 71.4%; Pred. No. 12;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLWFEID 7  
Db 390 FLWFEVD 396

##### RESULT 2

A10180 conserved hypothetical protein YP01485 [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004

C/Accession: A10180

R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; FMID:11586360

A/Accession: A10180

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-587 <KDR>

A/Cross-references: UNIPROT:Q8G34; GB:AL590842; PTDN:CAQ90308.1; PTD:G15979527; GSPDB:G

C/Genetics:

A/Gene: YP01485

C/Superfamily: uncharacterized conserved protein

Query Match 78.0%; Score 39; DB 2; Length 587;  
Best Local Similarity 85.7%; Pred. No. 15;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 FMEIDIV 9  
|||||  
Db 272 FMEIDIV 278

## RESULT 3

S40136  
T-cell receptor V-alpha 8.1b - human  
C:Species: Homo sapiens (man)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999  
C:Accession: S40136  
R:Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.  
submitted to the EMBL Data Library, February 1993  
A:Reference number: S40133  
A:Accession: S40136  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-110 <PLA>  
A:Cross-references: EMBL:X70308; NID:g437034; PIDN:CAA49787.1; PID:g437035  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: T-cell receptor

Query Match 74.0%; Score 37; DB 2; Length 110;  
Best Local Similarity 55.6%; Pred. No. 5.8; Indels 1; Gaps 0;  
Matches 5; Conservative 3; Mismatches 0; Gaps 0;

QY 1 FLMFIDIV 9  
|||||  
Db 10 FLMQLDLV 18

## RESULT 4

C29774  
T-cell receptor alpha chain precursor V region (HAP41) - human  
C:Species: Homo sapiens (man)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-May-1997  
C:Accession: C29774  
R:Yoshikai, Y.; Kimura, N.; Toyonaga, B.; Mak, T.W.  
J. Exp. Med. 164, 90-103, 1986  
A:Title: Sequences and repertoire of human T cell receptor alpha chain variable region 9  
A:Reference number: A92778; MUID:86253078; PMID:3088203  
A:Accession: C29774  
A:Molecule type: mRNA  
A:Residues: 1-134 <YOS>  
C:Genetics:  
A:Map position: 14q11.2  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: T-cell receptor

Query Match 74.0%; Score 37; DB 2; Length 134;  
Best Local Similarity 55.6%; Pred. No. 7.2; Indels 1; Gaps 0;  
Matches 5; Conservative 3; Mismatches 1; Gaps 0;

QY 1 FLMFIDIV 9  
|||||  
Db 10 FLMQLDLV 18

## RESULT 5

AG1585  
Hypothetical protein lin1224 [imported] - Listeria innocua (strain Clp11262)  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AG1585  
R:Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baguero, F.; Berche, P.; Bioecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fathi, H.  
Science 294, 849-852, 2001  
A:Authors: Krefit, U.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AG1585

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-150 <GLA>  
A:Cross-references: UNIPROT:Q92CE8; GB:AL592022; PIDN:CAC96455.1; PID:g16413698; GSPDB:GT  
C:Experimental source: strain Clp11262  
A:Gene: lin1224

Query Match 74.0%; Score 37; DB 2; Length 150;  
Best Local Similarity 77.8%; Pred. No. 8.1; Indels 2; Gaps 0;  
Matches 7; Conservative 0; Mismatches 2; Gaps 0;

QY 1 FLMFIDIV 9  
|||||  
Db 116 FKMFEIDKV 124

## RESULT 6

AH1231  
Hypothetical protein lmo1256 [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AH1231  
R:Glaser, P.; Frangul, L.; Buchrieser, C.; Amend, A.; Baguero, F.; Berche, P.; Bioecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fathi, H.  
Science 294, 849-852, 2001  
A:Authors: Krefit, U.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Ma  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AH1231  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-150 <GLA>  
A:Cross-references: UNIPROT:Q8V7L8; GB:NC\_003210; PIDN:CAC99334.1; PID:g16410672; GSPDB:GT  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo1256

Query Match 74.0%; Score 37; DB 2; Length 150;  
Best Local Similarity 77.8%; Pred. No. 8.1; Indels 2; Gaps 0;  
Matches 7; Conservative 0; Mismatches 2; Gaps 0;

QY 1 FLMFIDIV 9  
|||||  
Db 116 FKMFEIDKV 124

## RESULT 7

B86815  
Hypothetical protein ypgd [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: B86815  
R:Botolin, A.; Winkler, P.; Mauger, S.; Jailon, O.; Malarme, K.; Welschenbach, J.; Ehrli  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis 88  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Accession: B86815  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-529 <STO>  
A:Cross-references: UNIPROT:Q9CF66; GB:AE005176; PID:g12724521; PIDN:AAK05620.1; GSPDB:GT  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: ypgd

Query Match 74.0%; Score 37; DB 2; Length 529;

Best Local Similarity 66.7%; Pred. No. 31;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 FLMFEIDI 9  
Db 49 FLMFLADII 57

RESULT 8  
A:Accession: A86641  
hypotheoretical protein ybdc [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: A86641  
R:Botolin, A.; Wincker, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s8  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Status: preliminary  
A:Accession: A86641  
A:Residues: 1-269 <STO>  
A:Cross-references: UNIPROT:P75771; GB:AB005176; PID:g12722977; PIDN:AAK04227.1; GSPDB:C  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: ybdc  
C:Superfamily: stage III sporulation protein; stage III sporulation protein homology

Query Match 72.0%; Score 36; DB 2; Length 269;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLMFEI 6  
Db 155 FLMFEI 160

RESULT 9  
E64815  
Ybho protein - Escherichia coli (strain K-12)  
N:Alternate names: protein b0789  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C:Accession: E64815  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A.; Rose, D.U.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: E64815  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-413 <BLAT>  
A:Cross-references: UNIPROT:P75771; GB:AB000181; GB:U00096; NID:g1786998; PIDN:AAC73876.  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: ybho  
C:Superfamily: cardiolipin synthase

Query Match 72.0%; Score 36; DB 2; Length 413;  
Best Local Similarity 62.5%; Pred. No. 37;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEIDI 8  
Db 40 FLMFEIDV 47

RESULT 10  
D85587  
Probable synthetase ybho [imported] - Escherichia coli (strain O157:H7, substrain EDL933  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004

C:Accession: D85587  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grobeck, B.J.; Davis, N.W.; Lim, A.; Dialantia, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: D85587  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-413 <STO>  
A:Cross-references: UNIPROT:P75771; GB:AB005174; NID:g12513778; PIDN:AGS5160.1; GSPDB:G  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: ybho  
C:Superfamily: cardiolipin synthase

Query Match 72.0%; Score 36; DB 2; Length 413;  
Best Local Similarity 62.5%; Pred. No. 37;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEIDI 8  
Db 40 FLMFEIDV 47

RESULT 11  
C90737  
cardiolipin synthetase [imported] - Escherichia coli (strain O157:H7, substrain RMD 050  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: C90737  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno  
A:Reference number: A96629; MUID:21156231; PMID:11258796  
A:Accession: C90737  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-413 <NAV>  
A:Cross-references: UNIPROT:P75771; GB:BA000007; PIDN:BA834290.1; PID:g13360326; GSPDB:G  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: EC80867  
C:Superfamily: cardiolipin synthase

Query Match 72.0%; Score 36; DB 2; Length 413;  
Best Local Similarity 62.5%; Pred. No. 37;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEIDI 8  
Db 40 FLMFEIDV 47

RESULT 12  
D69457  
hypotheoretical protein AF161 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: D69457  
R:Klink, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
A.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gokeyne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Moese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: D69457  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-250 <KLE>

A;Cross-references: UNIPROT:O28612; GB:AE000989; GB:AE000782; NID:g2689312; PIDN:AA8961

Query Match 70.0%; Score 35; DB 2; Length 250;  
 Best Local Similarity 50.0%; Pred. No. 34;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LMFEID 8  
 Db 101 LMFEVD 108

RESULT 13  
 B82372  
 Spook-related protein VC0039 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)

C;Species: *Vibrio cholerae*  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C;Accession: B82372

R;Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.;  
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bae, S.; Qin, H.; Dragol, I.; Sellers, R.  
 L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MID:20406833; PMID:10952301

A;Accession: B82372

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-280 <HEI>

A;Cross-references: UNIPROT:Q9KVV0; GB:AE004096; GB:AE003852; NID:g9654440; PIDN:AA9321

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC0039

C;Superfamily: *Bacillus subtilis* hypothetical protein ysa1

Query Match 70.0%; Score 35; DB 2; Length 280;  
 Best Local Similarity 83.3%; Pred. No. 38;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LMFEID 7  
 Db 229 LMFEVD 234

RESULT 14

T08856

hypothetical protein A\_TM017A05.5 - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C;Accession: T08856

R;Waterston, R.

Submitted to the EMBL Data Library, October 1997

A;Reference number: Z16500

A;Accession: T08856

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-648 <WAT>

A;Cross-references: UNIPROT:Q95X14; EMBL:AF024504; NID:g2435510; PID:g2435516

A;Experimental source: cultivar Columbia

C;Genetics:

A;Map position: 4

A;Intons: 47/3; 90/1; 120/1; 185/1; 208/2; 222/3; 570/3

A;Note: A\_TM017A05.5

Query Match 70.0%; Score 35; DB 2; Length 648;  
 Best Local Similarity 83.3%; Pred. No. 93;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LMFEID 7  
 Db 175 LMFEVD 180

RESULT 15

T25325  
 hypothetical protein T26H2.7 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T25325

R;Matthews, L.

Submitted to the EMBL Data Library, November 1996

A;Reference number: Z20016

A;Accession: T25325

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-848 <MTL>

A;Cross-references: UNIPROT:O18139; EMBL:Z82055; PIDN:CA804848.1; GSPDB:GN00023; CESP:T26

A;Experimental source: clone T26H2

C;Genetics:

A;Gene: CESP:T26H2.7

A;Map position: 5

A;Intons: 41/2; 123/3; 176/1; 254/2; 312/1; 388/1; 425/3; 466/1; 514/3; 557/3; 628/3; 6

C;Superfamily: *Caenorhabditis elegans* hypothetical protein T26H2.7

Query Match 70.0%; Score 35; DB 2; Length 848;  
 Best Local Similarity 71.4%; Pred. No. 1,2e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 LMFEID 8  
 Db 699 LMFEVD 705

RESULT 16

T00842  
 probable histidine kinase [imported] - *Arabidopsis thaliana*

N;Alternate names: hypothetical protein T13L16.16

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004

C;Accession: T00842; G84556

R;de la Baudiere, M.; Hameed, A.; Gnoj, L.; Jensen, K.; Shohdy, N.; Gottesman, T.; Haberm

McCombie, W.R.

Submitted to the EMBL Data Library, January 1999

A;Description: A. thaliana BAC T13L16 from chromosome IV, top arm.

A;Reference number: Z14205

A;Accession: T00842

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1190 <DEL>

A;Cross-references: UNIPROT:Q7XJR4; EMBL:AC003952; NID:g2708736; PID:g2708752

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, B.

euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A;Reference number: A84420; MID:20083487; PMID:10617197

A;Accession: G84556

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1190 <STO>

A;Cross-references: GB:AE002093; NID:g2708752; PIDN:AA03576.1; GSPDB:GN00139

C;Genetics:

A;Gene: T13L16.16; A2917820

A;Map position: 2

A;Intons: 68/3; 149/3; 407/3; 490/2; 543/3; 586/1; 616/1; 704/2; 718/3; 764/3; 1

Query Match 70.0%; Score 35; DB 2; Length 1190;  
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LMFEID 7  
 Db 671 LMFEVD 676

RESULT 17

T52459  
sensory transduction histidine kinase 1 [validated] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 09-Jul-2004  
C:Accession: T52459  
R:Uro, T.; Yamaguchi-Shinozaki, K.; Hirayama, T.; Shinozaki, K.  
A:Title: A transmembrane hybrid-type histidine kinase in Arabidopsis functions as osm  
A:Reference number: 226082  
A:Accession: T52459  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1207 <URA>  
A:Cross-references: UNIPROT:Q9SXL4; EMBL:AB010914; PIDN:BAA32085.1  
C:Genetics:  
A:Gene: ATHK1  
A:Map position: 2  
C:Function:  
A:Description: involved in osmoregulation [validated, MUID:99418886]; probably functions  
C:Keywords: transmembrane protein

Query Match 70.0%; Score 35; DB 2; Length 1207;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFED 7  
Db 688 LWFED 693

RESULT 18  
A12244  
dihydropyrimidin aldolase [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C:Accession: A12244  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasa moto, S.; Watanabe, A.; Iriguchi,  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-121 <KUR>  
A:Cross-references: UNIPROT:Q8YRD6; GB:BA000019; PIDN:BAF75211.1; PID:917132645; GSPDB:C  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr3512  
C:Superfamily: dihydropyrimidin aldolase folA; dihydropyrimidin aldolase homology

Query Match 68.0%; Score 34; DB 2; Length 121;  
Best Local Similarity 66.7%; Pred. No. 24;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFEID 8  
Db 27 WFEID 32

RESULT 19  
AB0185  
probable membrane protein YPO1518 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AB0185  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchell, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
H. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AB0185  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-175 <KUR>  
A:Cross-references: UNIPROT:Q8ZG02; GB:AL590842; PIDN:CAC90341.1; PID:915979560; GSPDB:C  
C:Genetics:  
A:Gene: YPO1518

Query Match 68.0%; Score 34; DB 2; Length 175;  
Best Local Similarity 57.1%; Pred. No. 36;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LWFED 7  
Db 73 LWFED 79

RESULT 20  
F71612  
glycosyl transferase (novel euk. family) PFB0515w - malaria parasite (Plasmodium falciparum  
C:Species: Plasmodium falciparum  
C:Date: 13-Nov-1998 #sequence\_revision 13-Nov-1998 #text\_change 09-Jul-2004  
C:Accession: F71612  
R:Gardner, M.J.; Tetteh, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;  
Peters, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.  
Science 282, 1126-1132, 1998  
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.  
A:Reference number: A71600; MUID:99021743; PMID:9804551  
A:Accession: F71612  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-209 <GAR>  
A:Cross-references: UNIPROT:Q96196; GB:AE001400; GB:AE001362; NID:93845204; PIDN:AAC7189  
A:Experimental source: clone 3D7  
C:Genetics:  
A:Gene: PFB0515w

Query Match 68.0%; Score 34; DB 2; Length 209;  
Best Local Similarity 75.0%; Pred. No. 43;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LWFED 8  
Db 15 LWFED 22

RESULT 21  
AB1468  
ribosomal protein L1 [imported] - Listeria innocua (strain C11p11262)  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AB1468  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
D.; Dominguez-Bernal, G.; Duchaud, S.; Durand, L.; Dussurge, O.; Entian, K.D.; Fsihi, H.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schueter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AB1468  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-229 <GLA>  
A:Cross-references: UNIPROT:Q92R26; GB:AL592022; PIDN:CAC95514.1; PID:916412710; GSPDB:C  
A:Experimental source: strain C11p11262  
C:Genetics:  
A:Gene: rplA  
C:Superfamily: ribosomal protein L1p/L10e

Query Match 68.0%; Score 34; DB 2; Length 229;  
Best Local Similarity 57.1%; Pred. No. 47;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 3 WFEIDIV 9  
|||  
Db 108 WFEIDIV 114

## RESULT 22

C90024  
hypothetical protein [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: C90024  
R:Kutoda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; PMID:21311952; PMID:11418146  
A:Accession: C90024  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-254 <KUR>  
A:Cross-references: UNIPROT:Q99S08; GB:BA000018; P1D:G13702064; P1DN:BA843356.1; GSPDB:C  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: SA2059

Query Match 68.0%; Score 34; DB 2; Length 254;  
Best Local Similarity 71.4%; Pred. No. 53;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 3 WFEIDIV 9  
|||  
Db 191 WFEIDIV 197

## RESULT 23

T05592

tyrosine transaminase homolog F9D16.70 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T05592  
R:Byvan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hohenberg, J.; Mewes, H.W.; Mayer, K.F. submitted to the Protein Sequence Database, February 1999  
A:Reference number: Z15419  
A:Accession: T05592  
A:Molecule type: DNA  
A:Residues: 1-422 <BEV>  
A:Cross-references: UNIPROT:Q99S06; EMBL:AL035394  
A:Experimental source: cultivar Columbia; BAC clone F9D16  
C:Genetics:  
A:Map position: 4  
A:Introns: 84/2; 197/3; 270/3; 290/3; 342/3; 374/1  
A:Note: F9D16.70  
C:Superfamily: tyrosine transaminase

Query Match 68.0%; Score 34; DB 2; Length 422;  
Best Local Similarity 62.5%; Pred. No. 91;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLMFEIDI 8  
|||  
Db 340 FLMFEIDL 347

## RESULT 24

S44620

C50C3.1 protein - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S44620  
R:Favell, A.D. submitted to the EMBL Data Library, May 1993

A:Description: Sequence of the C. elegans cosmid C50C3.

A:Reference number: S44620  
A:Accession: S44620  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-463 <RAV>  
A:Cross-references: UNIPROT:P34366; EMBL:L14433; NID:G289649; P1D:G289653  
C:Genetics:  
A:Introns: 113/2; 125/3; 142/3; 181/1; 248/2; 355/3; 405/2

Query Match 68.0%; Score 34; DB 2; Length 469;  
Best Local Similarity 57.1%; Pred. No. 16+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLMFEID 7  
|||  
Db 67 FLMFEID 73

## RESULT 25

S55593

membrane protein SLS1 precursor - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein L3159; protein YLR139C  
C:Species: Saccharomyces cerevisiae  
C:Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004  
C:Accession: S55593; S59335; S64981; S64988; S72478  
R:Dufour, M.; Rouillard, J.; Mandart, E.; Lacroite, F. submitted to the EMBL Data Library, February 1995  
A:Description: SLS1, a new Saccharomyces cerevisiae gene, presenting a phenotype of synt  
A:Reference number: S55593  
A:Accession: S55593  
A:Molecule type: DNA  
A:Residues: 1-643 <DUF>  
A:Cross-references: UNIPROT:P42900; EMBL:Z48452; NID:G683706; P1D:G683707  
R:Delius, H.

submitted to the EMBL Data Library, June 1995  
A:Description: 36.8 kb of S.cerevisiae chromosome XII including ACE2, CK11, PDC5, SLS1,  
A:Reference number: S59313  
A:Accession: S59335  
A:Molecule type: DNA  
A:Residues: 1-643 <DEL>  
A:Cross-references: EMBL:X91258; NID:G995666; P1D:G995709  
A:Experimental source: strain S288C  
R:Delius, H.; Hebling, U. submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64967  
A:Accession: S64981  
A:Molecule type: DNA  
A:Residues: 1-643 <DEM>  
A:Cross-references: EMBL:Z73311; NID:G1360558; P1D:G245576; P1D:G1360559; MIPS:YLR139C  
A:Experimental source: strain S288C  
R:Legier, M.; Mueller-Auer, S.; Bruckner, M. submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64987  
A:Accession: S64988  
A:Molecule type: DNA  
A:Residues: 1-643 <RIE>  
A:Cross-references: EMBL:Z73311; NID:G1360558; P1D:G245576; P1D:G1360559; MIPS:YLR139C  
A:Experimental source: strain S288C  
R:Rouillard, J.M.; Dufour, M.E.; Theunissen, B.; Mandart, E.; Dujardin, G.; Lacroite, F.  
Mol. Gen. Genet. 252, 700-708, 1996  
A:Title: SLS1, a new Saccharomyces cerevisiae gene involved in mitochondrial metabolism.  
A:Reference number: S72478; PMID:97074886; PMID:8917313  
A:Accession: S72478  
A:Molecule type: DNA  
A:Residues: 1-643 <ROU>  
A:Cross-references: EMBL:Z48452; NID:G683706; P1DN:CA88362.1; P1D:G683707  
C:Genetics:  
A:Gene: SCD:SLS1  
A:Cross-references: SGD:S0004129; MIPS:YLR139C  
A:Map position: 12R  
A:Genome: nuclear  
C:Keywords: mitochondrion

Query Match 68.0%; Score 34; DB 2; Length 643;  
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFEBID 7  
 Db 477 LWFEBID 482

## RESULT 26

hypothetical protein BID4.140 [imported] - Neurospora crassa

C/Species: Neurospora crassa

C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

R/Schulte, U.; Aign, V.; Hohnel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, May 2000

A/Accession: T49399

A/Reference number: 225022

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-912 <SCH>

A/Cross-references: EMBL:AL355928, GSPDB:GN00116, NCSP:BID4.140

A/Experimental source: BAC clone BID4; strain OR74A

C/Genetics:

A/Gene: NCSP:BID4.140

A/Map position: 6

A/Introns: 702/1; 766/3

Qy 2 LWFEBID 8  
 Db 415 LWFEBID 421

## RESULT 27

probable pre-mRNA splicing factor - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C/Accession: T38841

R/Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1995

A/Reference number: 221815

A/Accession: T38841

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-2363 <GEN>

A/Cross-references: UNIPROT:O14187, EMBL:Z98530, PIDD:CA11062.1, GSPDB:GN00066, SPDB:SF

A/Experimental source: strain 972h-; cosmid c4F8

C/Genetics:

A/Map position: 1

Qy 1 LWFEBID 7  
 Db 948 LWFEBID 954

## RESULT 28

C82565

hypothetical protein XP2378 [imported] - Xylella fastidiosa (strain 9a5c)

C/Species: Xylella fastidiosa

C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen-

Nature 406, 151-157, 2000

A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A/Reference number: A82515; PMID:20365717; PMID:10910347

A/Note: for a complete list of authors see reference number A59328 below

A/Accession: C82565

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-55 <SIM>

A/Cross-references: UNIPROT:Q9PAW7, GB:AE004047, GB:AE003849, NID:99107548, PIDD:AAF6517

A/Experimental source: strain 9a5c

R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Melo, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitzajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E

A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A/Contents: annotation

A/Reference number: A59328

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

A/Contents: annotation

## RESULT 29

hypothetical protein XP2774 [imported] - Xylella fastidiosa (strain 9a5c)

C/Species: Xylella fastidiosa

C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C/Accession: A82519

R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen-

Nature 406, 151-157, 2000

A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A/Reference number: A82515; PMID:20365717; PMID:10910347

A/Note: for a complete list of authors see reference number A59328 below

A/Accession: A82519

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-126 <SIM>

A/Cross-references: UNIPROT:Q9P9U7, GB:AE004082, GB:AE003849, NID:99108003, PIDD:AAF6555

A/Experimental source: strain 9a5c

R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Melo, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitzajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E

A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

Query Match 66.0%; Score 33; DB 2; Length 126;  
 Best Local Similarity 83.3%; Pred. No. 39;

Qy 1 LWFEBID 7  
 Db 24 LWFEBID 30

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLMFBI 6  
 Db 20 FLMFBI 25

## RESULT 30

hypotheical protein homolog lin0633 [imported] - *Listeria innocua* (strain Clp11262)  
 C/Species: *Listeria innocua*  
 C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: A11511

R/Glaser, P.; Frangoul, L.; Bucheleier, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Ertan, K.D.; Feinl, H.; Jones, L.M.; Kars, U.  
 Science 294, 849-852, 2001  
 A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kuzupat, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.  
 A/Reference number: AB1077; PMID:21537279; PMID:11679669  
 A/Accession: A11511  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-143 <GLA>  
 A/Cross-references: UNIPROT:Q92E28; GB:AL592022; PIDN:CAC95865.1; PID:g16413073; GSPDB:G  
 A/Experimental source: strain Clp11262  
 C/Genetics:  
 A/Gene: lin0633  
 C/Superfamily: *Bacillus subtilis* conserved hypothetical protein ydgb

Query Match 66.0%; Score 33; DB 2; Length 143;  
 Best Local Similarity 57.1%; Pred. No. 44;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFBI 7  
 Db 66 YLMFHD 72

## RESULT 31

dhHydronopter aldolase (EC 4.1.2.25) - *Synechocystis* sp. (strain PCC 6803)  
 C/Species: *Synechocystis* sp.  
 A/Variety: PCC 6803  
 C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
 C/Accession: S76177

R/Kane, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asami, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996  
 A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.  
 A/Reference number: S74322; PMID:97061201; PMID:8905231  
 A/Accession: S76177  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-150 <KAN>

A/Cross-references: EMBL:D90914; GB:AB001339; NID:g1653477; PIDN:BA18436.1; PID:g165352  
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C/Superfamily: dhHydronopter aldolase fold; dhHydronopter aldolase homology  
 C/Keywords: aldehyde-lyase; carbon-carbon lyase; folate biosynthesis  
 F.34-150/Domain: dhHydronopter aldolase homology <DNA>

Query Match 66.0%; Score 33; DB 1; Length 150;  
 Best Local Similarity 66.7%; Pred. No. 47;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 WFEIDI 8  
 Db 59 WFEVDL 64

## RESULT 32

A82343  
 conserved hypothetical protein VC0266 [imported] - *Vibrio cholerae* (strain N16961 seroGrC  
 C/Species: *Vibrio cholerae*  
 C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C/Accession: A82343

R/Heidelberg, U.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P. 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A/Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 A/Reference number: A82035; PMID:120406833; PMID:10952301  
 A/Accession: A82343  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-190 <HEI>

A/Cross-references: UNIPROT:Q9KV90; GB:AE004116; GB:AE003852; NID:g9654687; PIDN:AAF3441  
 A/Experimental source: serogroup O1, strain N16961, biotype El Tor  
 C/Genetics:  
 A/Gene: VC0266  
 A/Map position: 1

Query Match 66.0%; Score 33; DB 2; Length 190;  
 Best Local Similarity 85.7%; Pred. No. 60;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 WFEIDI 9  
 Db 172 WFEQDIV 178

## RESULT 33

hypotheical protein CPJ0978 [imported] - *Chlamydia pneumoniae* (strain J138)  
 C/Species: *Chlamydia pneumoniae*, *Chlamydia pneumoniae*  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001

R/Shital, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is Nucleic Acids Res. 28, 2311-2314, 2000  
 A/Title: Comparison of whole genome sequences of *Chlamydia pneumoniae* J138.  
 A/Reference number: A86491; PMID:20330349; PMID:10871362  
 A/Accession: G86612  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-218 <STO>

A/Cross-references: GB:BA000008; NID:g8979351; PIDN:BA99185.1; GSPDB:GNO0142  
 A/Experimental source: strain J138  
 C/Genetics:  
 A/Gene: CPJ0978

Query Match 66.0%; Score 33; DB 2; Length 218;  
 Best Local Similarity 71.4%; Pred. No. 70;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFBI 7  
 Db 120 FLMFBI 126

## RESULT 34

protein T10022.3 [imported] - *Arabidopsis thaliana*  
 C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C/Accession: A86316

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.H.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K., ansen, N.F.; Hughes, B.; Huzar, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.



A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: A86316  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-226 <STO>  
A:Cross-references: UNIPROT:Q9LM40; GB:AE005172; NID:98671786; PIDN:AAF8392.1; GSPDB:GN  
C:Genetics:  
A:Gene: T10022.3  
A:Map position: 1

Query Match 66.0%; Score 33; DB 2; Length 226;  
Best Local Similarity 71.4%; Pred. No. 72;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFIDIV 7  
Db 196 FLMFIDIV 202

RESULT 35  
H82616  
hypothetical protein XFI958 [imported] - *Xylella fastidiosa* (strain 9asc)  
C:Species: *Xylella fastidiosa*  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: H82616  
R:anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: H82616  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-269 <STM>  
A:Cross-references: UNIPROT:Q9PC27; GB:AE004015; GB:AE003849; NID:99107057; PIDN:AAF8476  
A:Experimental source: strain 9asc  
R:Simpon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Asencio, M.; Alvarenga, R.; A  
Brites, M.R.S.; Bueno, M.R.P.; Canarso, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carver, H  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Petro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm  
J.D.; Jungueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Klieger, J.B.; Kuramae, E.E.; Laig  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sasaki  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tshako, M.H.; Valada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XFI958

Query Match 66.0%; Score 33; DB 2; Length 269;  
Best Local Similarity 62.5%; Pred. No. 87;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMFIDIV 9  
Db 242 LMFIDIV 249

RESULT 36  
D95396  
probable polyketide synthase protein [imported] - *Sinorhizobium meliloti* (strain 1021)  
C:Species: *Sinorhizobium meliloti*  
C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C:Accession: D95396  
R:Barnett, M.J.; Fisher, R.P.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe  
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*

A:Reference number: A95262; MUID:21396509; PMID:11481432  
A:Accession: D95396  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-269 <KUR>  
A:Cross-references: UNIPROT:Q9Y112; GB:AE006469; PIDN:AAK5734.1; PID:914524229; GSPDB:G  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Galibert, F.; Fijan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chaim, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: Sma1963  
A:Genome: plasmid

Query Match 66.0%; Score 33; DB 2; Length 269;  
Best Local Similarity 66.7%; Pred. No. 87;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFIIDIV 8  
Db 107 WFIIDIV 112

RESULT 37  
B71876  
probable prolipoprotein diacylglycerol transferase - *Helicobacter pylori* (strain J99)  
C:Species: *Helicobacter pylori*  
A:Variety: strain J99  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: B71876  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dojic, P.C.; Smith, D.R.;  
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Voyis, G.F.;  
Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A:Reference number: A71800; MUID:99120557; PMID:9923682  
A:Accession: B71876  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-283 <ARN>  
A:Cross-references: UNIPROT:Q9ZKP6; GB:AE001518; GB:AE001439; NID:94155454; PIDN:AAD0647  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: lgt  
C:Superfamily: prolipoprotein diacylglycerol transferase

Query Match 66.0%; Score 33; DB 2; Length 283;  
Best Local Similarity 66.7%; Pred. No. 92;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLMFIDIV 9  
Db 63 FLMFIDIV 71

RESULT 38  
C64639  
prolipoprotein diacylglycerol transferase (EC 2.4.99.-) - *Helicobacter pylori* (strain 26.  
C:Species: *Helicobacter pylori*  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: C64639  
R:Tombs, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey, L.  
Nature 389, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.N.  
A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A:Reference number: A64520; MUID:97394467; PMID:9252185

A;Accession: G64639  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-284 <TOM>  
 A;Cross-references: UNIPROT:Q25609; GB:AE000604; GB:AE000511; NID:g2314086; PIDN:AAD0799  
 C;Superfamily: Protoprotein diacylglycerol transferase  
 C;Keywords: glycosyltransferase

Query Match 66.0%; Score 33; DB 1; Length 284;  
 Best Local Similarity 66.7%; Pred. No. 93;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 FLMFEDIV 9  
 Db 63 FLMAELGIV 71

## RESULT 39

AG2429  
 delta-9 desaturase [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120  
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
 C;Accession: AG2429  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
 A;Reference number: AB1807; MUID:21595285; PMID:11759840  
 A;Accession: AG2429  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-285 <KUR>  
 A;Cross-references: UNIPROT:Q8YME3; GB:BA000019; PIDN:BA076690.1; PID:g17134129; GSPDB:G  
 A;Experimental source: strain PCC 7120  
 C;Genetics:  
 A;Gene: desC  
 C;Superfamily: delta-9 acyl-CoA desaturase; stearyl-CoA desaturase homology

Query Match 66.0%; Score 33; DB 2; Length 285;  
 Best Local Similarity 50.0%; Pred. No. 93;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLMFEDIV 8  
 Db 250 FQWMEVDV 257

## RESULT 40

F72011

hypothetical protein CP0878 [imported] - Chlamydia pneumoniae (strains CWL029 and AR  
 C;Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
 C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C;Accession: F72011; H81528

R;Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
 Nature Genet. 21, 385-389, 1999  
 A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
 A;Reference number: A72000; MUID:99206606; PMID:10192388

A;Accession: F72011  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-343 <ARN>

A;Cross-references: UNIPROT:Q9Z6T1; GB:AE001678; GB:AE001663; NID:g4377301; PIDN:AAD1911

A;Experimental source: strain CWL029  
 R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolony, J.; McClarty, G.; Salzberg,  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A;Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.  
 A;Reference number: A81500; MUID:20150255; PMID:10684935

A;Accession: H81528  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-343 <REA>

A;Cross-references: GB:AE002246; GB:AE002161; NID:g7189785; PIDN:AA038666.1; PID:g718979  
 A;Experimental source: strain AR39, HL cells  
 C;Genetics:  
 A;Gene: CP0978; CP0878  
 C;Superfamily: Chlamydia hypothetical protein CP0978

Query Match 66.0%; Score 33; DB 2; Length 343;  
 Best Local Similarity 71.4%; Pred. No. 11e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLMFEDIV 7  
 Db 120 FLMFEDIV 126

Search completed: January 12, 2005, 20:15:46  
 Job time : 19.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:38:08 ; Search time 85 Seconds  
(without alignments)  
60.922 Million cell updates/sec

Title: US-09-870-216C-7  
Perfect score: 50  
Sequence: 1 FLWFEIDIV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	78.0	266	2 Q6D9R7	Q6D9R7 erwina car
2	39	78.0	474	2 Q14192	Q14192 schizosacch
3	39	78.0	587	2 Q8ZG34	Q8ZG34 yersinia pe
4	39	78.0	587	2 AAS61616	AAS61616 yersinia
5	39	78.0	596	2 Q81IA6	Q81IA6 plasmidium
6	39	78.0	1714	2 Q61Z03	Q61Z03 oestrinia nu
7	39	78.0	1714	2 AAT37678	AAT37678 oestrinia
8	38	76.0	588	2 Q7MZS2	Q7MZS2 photorhabdu
9	37	74.0	150	2 Q92CE8	Q92CE8 listeria in
10	37	74.0	150	2 Q8Y7L8	Q8Y7L8 listeria mo
11	37	74.0	198	2 Q7R8S9	Q7R8S9 plasmidium
12	37	74.0	322	2 Q7UJ06	Q7UJ06 rhodospirell
13	37	74.0	529	2 Q9CFE6	Q9CFE6 lactococcus
14	37	74.0	586	2 Q7N267	Q7N267 photorhabdu
15	37	74.0	713	2 Q7RV54	Q7RV54 neurospora
16	37	74.0	731	2 Q74693	Q74693 neurospora
17	36	72.0	129	2 Q8DKV4	Q8DKV4 synecococc
18	36	72.0	264	2 Q8DD89	Q8DD89 vibrio vuln
19	36	72.0	266	2 Q7MGKO	Q7MGKO vibrio vuln
20	36	72.0	269	1 QXAL LACUA	Q9CJ72 lactococcus
21	36	72.0	270	2 Q87KCG	Q87KCG vibrio para
22	36	72.0	331	2 Q93X27	Q93X27 populus eur
23	36	72.0	413	1 YBHO_ECOLI6	Q8Jfj9 escherichia
24	36	72.0	413	1 YBHO_ECOLI	P75771 escherichia
25	36	72.0	445	2 Q70X56	Q70X56 scutigera c
26	36	72.0	445	2 CAD45022	CAD45022 scutigera
27	36	72.0	590	1 SYD PHOIL	Q7MB41 photorhabdu
28	36	72.0	700	1 HPSI_HUMAN	Q92902 homo sapien
29	36	72.0	700	2 Q8WKE5	Q8WKE5 homo sapien
30	36	72.0	704	1 HPSI_MOUSE	Q08983 mus musculu
31	36	72.0	706	2 Q99MK7	Q99MK7 rattus norv

32	35	70.0	76	2 Q935T6	Q935T6 streptococc
33	35	70.0	112	2 Q9WU87	Q9WU87 mus musculu
34	35	70.0	194	2 Q8VEN8	Q8VEN8 mus musculu
35	35	70.0	205	2 Q83A27	Q83A27 coxiella bu
36	35	70.0	232	2 Q81XE0	Q81XE0 homo sapien
37	35	70.0	250	2 Y661_ARCFU	Q28612 archaeoglob
38	35	70.0	261	2 Q8ES40	Q8ES40 oceanobactil
39	35	70.0	280	2 Q9KVV0	Q9KVV0 vibrio chol
40	35	70.0	285	1 FPG CORER	Q8FD17 corynebacte
41	35	70.0	309	2 Q7TRM3	Q7TRM3 mus musculu
42	35	70.0	309	2 Q8VFT9	Q8VFT9 mus musculu
43	35	70.0	311	2 Q7TRM1	Q7TRM1 mus musculu
44	35	70.0	311	2 Q7TRM2	Q7TRM2 mus musculu
45	35	70.0	311	2 Q8VF70	Q8VF70 mus musculu

## ALIGNMENTS

RESULT 1					
ID	Q6D9R7	PRELIMINARY:	PRT:	266 AA.	
AC	Q6D9R7				
DT	01-OCT-2004 (TrEMBLrel. 28, Created)				
DT	01-OCT-2004 (TrEMBLrel. 28, Last sequence update)				
DT	01-OCT-2004 (TrEMBLrel. 28, Last annotation update)				
DE	Putative plasmid transfer protein.				
CN	Name=trab; ORFNames=ECa0548;				
OS	Erwinia carotovora subsp. atroseptica SCRI1043.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Pectobacterium.				
OX	NCBI_TaxID=218491;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=SCRI1043;				
RA	Bell K.S., Sebahia M., Pritchard L., Holden M., Hyman L.J.,				
RA	Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,				
RA	Atkin R., Baason N., Brooks K., Chillingworth T., Clark K., Doggett J.,				
RA	Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,				
RA	Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,				
RA	Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.;				
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BX950851; CAG73463.1; -				
SQ	SEQUENCE 266 AA; 30560 MW; 8D7628F313223F8E CRC64;				
Query Match					
	Best Local Similarity	78.0%;	Score 39;	DB 2;	Length 266;
	Matches	6;	Conservative	1;	Mismatches
				0;	Indels
					Gaps
					0;
QY	2 LMFEIDI 8				
DB	134 LMFEIDV 140				
RESULT 2					
ID	Q14192	PRELIMINARY:	PRT:	474 AA.	
AC	Q14192				
DT	01-JAN-1999 (TrEMBLrel. 09, Created)				
DT	01-JAN-1999 (TrEMBLrel. 09, Last sequence update)				
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)				
DE	SPAC5684.03 protein (EC 2.6.1.-).				
CN	Name=SPAC5684.03;				
OS	Schizosaccharomyces pombe (Fission yeast).				
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;				
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;				
OX	NCBI_TaxID=4896;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=972h-;				
RC	MEDLINE=21848401; PubMed=118593360;				
RX	Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,				

RA Sgouros J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,  
 RA Woodward J., Volkstraet G., Aert R., Robben J., Gymnopoulos B.,  
 RA Welljens I., Vanstreels B., Rieger M., Schaefer M., Muller-Auer S.,  
 RA Gabell C., Fuchs M., Dusterhoft A., Fritzc C., Holzer E., Moesli D.,  
 RA Hilbert H., Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R.,  
 RA Pohl T.M., Egger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreaon S., Gloux S., Jelaire V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Ceretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussey D., Barrell B.G., Nurse P.,  
 RA "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 DR EMBL; 299261; CAB16394.1; -.  
 DR PIR; T38905; T38905.  
 DR Genedb\_Spmbe; SPAC5654.03; -.  
 DR GO; GO:0008483; F:transaminase activity; IEA.  
 DR GO; GO:0016740; F:transaminase activity; IEA.  
 DR InterPro; IPR002052; N6\_Mtase.  
 DR PROSITE; PS00092; N6\_MTASE; UNKNOWN\_1.  
 KW Amino transferase; Transferase.  
 SQ SEQUENCE 474 AA; 5316 MW; B546EABC7B884623 CRC64;

Query Match 78.0%; Score 39; DB 2; Length 474;  
 Best Local Similarity 71.4%; Pred. NO. 74;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FIMFED 7  
 Db 390 FIMFEDV 396

RESULT 3  
 ID Q8ZG34 PRELIMINARY; PRT; 587 AA.  
 AC Q8ZG34; Q7AVD6; Q7CHP0;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein YP01485  
 GN OrderedLocustNames=YP1375, YP01485, Y2684;  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Yersinia.  
 OX NCB1\_TaxID=632;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92 / Biovar Orientalis;  
 RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;  
 RA Parkhill J., Wren B.W., Thomson N.R., Titchall R.W., Holden M.T.G.,  
 RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Fellwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,  
 RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,  
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.,  
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."  
 RL Nature 413:523-527(2001).  
 RN  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=KIMS / Biovar Mediaevalis;

RX MEDLINE=22137863; PubMed=12142430;  
 RA Deng W., Burland V., Plunkett G., III, Boutin A., Mayhew G.F., Liss P.,  
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
 RA Retherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
 RA Perry R.D.,  
 RT "Genome sequence of Yersinia pestis KIM."  
 RL J. Bacteriol. 184:4601-4611(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=91001 / Biovar Mediaevalis;  
 RA Song Y., Tong Z., Wang L., Han Y., Qin H., Wang J., Li S., Guo Z.,  
 RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,  
 RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,  
 RA Yang R.,  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; A7414148; CAC90308.1; -.  
 DR EMBL; AB013870; AAM86237.1; -.  
 DR EMBL; AB017132; AAS61616.1; -.  
 DR PIR; A10180; A10180.  
 DR InterPro; IPR010272; DUF879.  
 DR Pfam; PF05947; DUF879; 1.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 587 AA; 67418 MW; 389A200B93753764 CRC64;

Query Match 78.0%; Score 39; DB 2; Length 587;  
 Best Local Similarity 85.7%; Pred. NO. 91;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFEIDIV 9  
 Db 272 WFEIDIV 278

RESULT 4  
 ID AAS61616 PRELIMINARY; PRT; 587 AA.  
 AC AAS61616;  
 DT 24-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 24-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN YP1375.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Yersinia.  
 OX NCB1\_TaxID=632;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=91001 / Biovar Mediaevalis;  
 RA Song Y., Tong Z., Wang L., Han Y., Qin H., Wang J., Li S., Guo Z.,  
 RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,  
 RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,  
 RA Yang R.,  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB017132; AAS61616.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 587 AA; 67418 MW; 389A200B93753764 CRC64;

Query Match 78.0%; Score 39; DB 2; Length 587;  
 Best Local Similarity 85.7%; Pred. NO. 91;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WFEIDIV 9  
 Db 272 WFEIDIV 278

RESULT 5  
 ID Q81IA6 PRELIMINARY; PRT; 596 AA.  
 AC Q81IA6;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=PF11\_0268;  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22255705; PubMed=12368864;  
 RA Gardner M.J., Hall N., Fung B., White O., Berriman M., Hyman R.W.,  
 RA Carlton J.M., Pain A., Nelson K.R., Bowman S., Paulsen I.T., James K.,  
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,  
 RA Chan M.S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,  
 RA Perceva M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,  
 RA Martin D.M., Faulkner A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,  
 RA Madden G.I., Cummings L.M., Subramanian G.M., Mungall C., Davis R.W.,  
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,  
 RA Fraser C.M., Barrell B.;  
 RT "Genome sequence of the human malaria parasite Plasmodium  
 RT falciparum."  
 RL Nature 419:498-511(2002).  
 DR EMBL; AB014839; AAN35852.1; -;  
 DR InterPro; IPR011043; Gal\_oxid\_central.  
 DR InterPro; IPR011498; Kelch\_2.  
 DR InterPro; IPR006652; Kelch\_rep.  
 DR Pfam; PF01344; Kelch\_1; 2.  
 DR Pfam; PF07646; Kelch\_2; 1.  
 KM Hypothetical protein\_2;  
 SQ SEQUENCE 596 AA; 68046 MW; B07DC58441A754AF CRC64;

Query Match 78.0%; Score 39; DB 2; Length 596;  
 Best Local Similarity 75.0%; Pred. No. 92;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLWFEID 8  
 Db 433 FSWFEVDI 440

RESULT 6  
 ID 061203 PRELIMINARY; PRT; 1714 AA.  
 AC 061203;  
 DT 05-JUL-2004 (TReMBLrel. 27, Created)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
 DE Cadherin A1.  
 OS Oestrinia nubilalis (European corn borer).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Pyraloidea;  
 OC Pyralidae; Pyraustinae; Ostrinia.  
 OX NCBI\_TaxID=29057;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Coates B.S., Sumerford D.V., Hellmich R.L.;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 11 cadherin domains.  
 DR EMBL; AY612336; AAT37678.1; -;  
 DR InterPro; IPR002126; Cadherin.  
 DR Pfam; PF00028; Cadherin; 5.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 10.  
 DR PROSITE; PS00232; CADHERIN\_1; 2.  
 DR PROSITE; PS50268; CADHERIN\_2; 11.  
 KM Calcium; Calcium-binding.  
 SQ SEQUENCE 1714 AA; 191900 MW; 89E3CAE0614654DE CRC64;

Query Match 78.0%; Score 39; DB 2; Length 1714;  
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 FLWFEID 7

Db 650 YLWFEID 656  
 RESULT 7  
 ID AAT37678 PRELIMINARY; PRT; 1714 AA.  
 AC AAT37678;  
 DT 01-JUN-2004 (TReMBLrel. 27, Created)  
 DT 01-JUN-2004 (TReMBLrel. 27, Last sequence update)  
 DT 01-JUN-2004 (TReMBLrel. 27, Last annotation update)  
 DE Cadherin A1.  
 OS Oestrinia nubilalis (European corn borer).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Pyraloidea;  
 OC Pyralidae; Pyraustinae; Ostrinia.  
 OX NCBI\_TaxID=29057;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Coates B.S., Sumerford D.V., Hellmich R.L.;  
 RT "Single nucleotide polymorphism (SNP) detection in proximity of two  
 RT putative toxin-binding regions of the candidate bacillus thuringiensis  
 RT resistance gene, cadherin, of Oestrinia nubilalis."  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY612336; AAT37678.1; -;  
 SQ SEQUENCE 1714 AA; 191900 MW; 89E3CAE0614654DE CRC64;

Query Match 78.0%; Score 39; DB 2; Length 1714;  
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLWFEID 7  
 Db 650 YLWFEID 656

RESULT 8  
 ID 07M2S2 PRELIMINARY; PRT; 588 AA.  
 AC 07M2S2;  
 DT 01-MAR-2004 (TReMBLrel. 26, Created)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Similar to unknown protein.  
 GN OrderedLocustNames=plu4202;  
 OS Photobacterium luminescens (subsp. laumondii).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Photobacterus.  
 OX NCBI\_TaxID=141679;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TT01;  
 RX MEDLINE=22957627; PubMed=14528314;  
 RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,  
 RA Taourit S., Bocs S., Bouraux-Eude C., Chaudier M., Charles J.-F.,  
 RA Daes A., Derose R., Derzelle S., Freysinet G., Gaudreau S.,  
 RA Médigue C., Lanois A., Powell K., Siguer P., Vincent R., Wingate V.,  
 RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.,  
 RT "The genome sequence of the entomopathogenic bacterium Photobacterus  
 RT luminescens."  
 RL Nat. Biotechnol. 21:1307-1313(2003).  
 DR EMBL; BX571873; CAE16574.1; -;  
 DR Photobact; plu4202; -;  
 DR InterPro; IPR010272; DUF679.  
 DR Pfam; PF05947; DUF679; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 588 AA; 67517 MW; F5D33F2A6A5DFD5A CRC64;

Query Match 76.0%; Score 38; DB 2; Length 588;  
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3 WFEIDIV 9

Db 273 WFEIDIV 279

## RESULT 9

092CE8 PRELIMINARY; PRT; 150 AA.

AC 092CE8 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Lm1224 protein.  
 GN OrderedLocusNames=lm1224;  
 OS *Listeria innocua*.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 NCBI\_TaxID=1642;  
 RX STRAIN=CLIP 11262 / Serovar 6a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 Charbit A., Chetouani F., Couve E., de Darvar A., Dehoux P.,  
 Entian K.-D., Dominguez-Bernal G., Duchaud E., Durant L., Duseigneur O.,  
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,  
 Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjati H.,  
 Norislet G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,  
 Vazquez-Boland J.-A., Voss H., Wehland J., Coesart P.;  
 RT "Comparative genomics of *Listeria* species."  
 RL Science 294:849-852(2001).  
 CC -1- SIMILARITY: Belongs to the Nudix hydrolase family.  
 DR EMBL AL596168; CAC96455.1; -  
 DR PIR AG1585; AG1585.  
 DR Listerlist; Lm1224; -  
 DR InterPro; IPR000086; NUDIX\_hydrolase.  
 DR Pfam; PF00293; NUDIX; 1.  
 DR PROSITE; PS00893; NUDIX; 1.  
 DR Complete proteome; Hydrolase.  
 SQ SEQUENCE 150 AA; 17636 MW; 205F376076CB8CE7 CRC64;

Query Match 74.0%; Score 37; DB 2; Length 150;  
 Best Local Similarity 77.8%; Pred. NO. 56;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLMFEIDIV 9  
 Db 116 FKMFEIDKV 124

RESULT 10  
 08Y7L8 PRELIMINARY; PRT; 150 AA.

AC 08Y7L8 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Lmo1256 protein.  
 GN OrderedLocusNames=lmo1256;  
 OS *Listeria monocytogenes*.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 NCBI\_TaxID=1639;  
 RX STRAIN=EGD-e / Serovar 1/2a;  
 RX MEDLINE=21537279; PubMed=11679669;  
 RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 Charbit A., Chetouani F., Couve E., de Darvar A., Denoux P.,  
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duseigneur O.,  
 Entian K.-D., Feihl H., Garcia-del Portillo F., Garrido P.,

RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,  
 Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjati H.,  
 Nordislet G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,  
 Vazquez-Boland J.-A., Voss H., Wehland J., Coesart P.;  
 RT "Comparative genomics of *Listeria* species."  
 RL Science 294:849-852(2001).  
 CC -1- SIMILARITY: Belongs to the Nudix hydrolase family.  
 DR EMBL AL591978; CAC99334.1; -  
 DR PIR AH1231; AH1231.  
 DR Listerlist; Lm1256; -  
 DR InterPro; IPR000086; NUDIX\_hydrolase.  
 DR Pfam; PF00293; NUDIX; 1.  
 DR PROSITE; PS00893; NUDIX; 1.  
 DR Complete proteome; Hydrolase.  
 SQ SEQUENCE 150 AA; 17795 MW; D72C7098952585B0 CRC64;

Query Match 74.0%; Score 37; DB 2; Length 150;  
 Best Local Similarity 77.8%; Pred. NO. 56;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FLMFEIDIV 9  
 Db 116 FKMFEIDKV 124

## RESULT 11

07RSS9 PRELIMINARY; PRT; 198 AA.

AC 07RSS9 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN Name=PY00274;  
 OS *Plasmodium yoelii yoelii*.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 NCBI\_TaxID=73239;  
 RX STRAIN=17XN1;  
 RX PubMed=13368865;  
 RA Carlson J.M., Angiolini S.V., Suh B.B., Kocif T.W., Petrea M.,  
 Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,  
 Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,  
 Shallow S.J., van Aken S.B., Riedmiller S.B., Feldblum T.V.,  
 Cho J.K., Quackenbush J., Sedegh M., Shoab A., Cummings L.M.,  
 Florens L., Yates F.R. III, Raine J.D., Sindén R.E., Harris M.A.,  
 Cunningham D.A., Pfeister P.R., Bergman L.W., Valdivia A.B.,  
 van Lin L.H., Jansz C.J., Waters A.P., Smith H.O., White O.R.,  
 Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,  
 Carucci D.J.;  
 RT "Genome sequence and comparative analysis of the model rodent malaria  
 parasite *Plasmodium yoelii yoelii*.";  
 RL Nature 419:512-519(2002).  
 CC -1- CAUTION: The sequence shown here is derived from an  
 EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is  
 preliminary data.  
 DR EMBL; AABJ0100077; EAA22282.1; -  
 DR Hypothetical protein.  
 SQ SEQUENCE 198 AA; 22441 MW; 50A2D7E56AB5E236 CRC64;

Query Match 74.0%; Score 37; DB 2; Length 198;  
 Best Local Similarity 75.0%; Pred. NO. 74;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLMFEIDIV 8  
 Db 33 FSWFEIDI 40

## RESULT 12

```

07U06
ID Q7U06 PRELIMINARY; PRT; 322 AA.
AC Q7U06;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=RB12213;
OC Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
RA Schleutner H., Mann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
KM EMBL; BX294154; CAD77454.1; -.
SQ SEQUENCE 322 AA; 35397 MW; 9790D69F2C0EB08B CRC64;

Query Match
Best Local Similarity 74.0%; Score 37; DB 2; Length 322;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLWFEID 7
Db 79 FIFWFDID 85

RESULT 13
ID Q9CF6 PRELIMINARY; PRT; 529 AA.
AC Q9CF6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE ABC transporter ATP binding and permease protein.
GN Name=YP9D; OrderedLocustNames=JL1522;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=114403;
RX MEDLINE=21235186; PubMed=11337471;
RA Boletín A., Wincker P., Manger S., Jallion O., Malarme K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
EMBL; AB006383; AA05620.1; -.
PIR; B86815; B86815.
DR GO; GO:0016021; C: integral to membrane, IEA.
DR GO; GO:0005524; F: ATP binding, IEA.
DR GO; GO:0004009; F: ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0001166; F: nucleotide binding, IEA.
DR GO; GO:0006810; P: transport, IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp.
DR Pfam; PF00664; ABC_membrane_1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transp.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS50929; ABC_TM1F; 1.

```

```

DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
SQ SEQUENCE 529 AA; 59763 MW; 7838906A4D09E0C5 CRC64;

Query Match
Best Local Similarity 74.0%; Score 37; DB 2; Length 529;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9
Db 49 FLWLAIDII 57

RESULT 14
ID Q7N267 PRELIMINARY; PRT; 586 AA.
AC Q7N267;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Similar to unknown protein.
GN OrderedLocustNames=plu3225;
OS Photobacterium luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photobacterus.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Tauchard E., Rusnlok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Bouraux-Eude C., Chandler M., Charles J.-F.,
RA Daasa B., Derose R., Derzelle S., Freysinet G., Gaudriault S.,
RA Médigue C., Lanois A., Powell K., Signier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photobacterus
RT luminescens."
RL Nat. Biotechnol. 21:1307-1313(2003).
DR EMBL; BX571869; CAE15599.1; -.
DR Photobact; plu3225; -.
DR InterPro; IPR010272; DUF879.
DR Pfam; PF05947; DUF879; 1.
KM Complete proteome.
SQ SEQUENCE 586 AA; 67513 MW; AA8AB383C031F30F CRC64;

Query Match
Best Local Similarity 74.0%; Score 37; DB 2; Length 586;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 WFEIDIV 9
Db 272 WFEIDIV 278

RESULT 15
ID Q7RV54 PRELIMINARY; PRT; 713 AA.
AC Q7RV54;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein (AP088906) clock-controlled gene-9
DE protein).
GN Name=NCU09559.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,

```

RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Janakiev P., Pedersen D., Nelson M., Washburne M.,  
 RA Selltreunikoit C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
 RA Koche G.O., Jedd G., Mewes W., Staben C., Marcotte B., Greenberg D.,  
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gierre S.,  
 RA Kanal M., Kamyselle M., Mauceli E., Bielke C., Rudd S., Frihman D.,  
 RA Krysstova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroen S.,  
 RA Cosoni C., Macino G., Catcheside D., Li W., Pratt R.J., Omani S.A.,  
 RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
 RA Varden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
 RA Narevig D.O., Alex L.A., Mannheim G., Ebbola D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,  
 RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa.",  
 RL Nature 0:0-0(2003).  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL: AABX0100379; EAA30248.1; -.  
 DR GO: GO:0009058; P:biomynthesis; IEA.  
 DR InterPro: IPR001296; Glyco\_transf\_1.  
 DR Pfam: PF00534; Glyco\_transf\_1; 1.  
 DR Hypothetical protein.  
 KW SEQUENCE 713 AA; 79717 MW; DA651BA0E524DBEC CRC64;

Query Match 74.0%; Score 37; DB 2; Length 713;  
 Best Local Similarity 75.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMFEDIV 9  
 DB 124 LMLEVDIV 131

RESULT 16  
 074693 PRELIMINARY; PRT; 731 AA.  
 ID 074693; 001308;  
 AC 074693; 001308;  
 DT 01-NOV-1998 (TRENBLrel. 08, Created)  
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Clock-controlled gene-9 protein.  
 GN Name:ccg-9;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97075124; PubMed=8917550;  
 RA Bell-Pedersen D., Shinohara M.L., Loros J.J., Dunlap J.C.;  
 RT "Circadian clock-controlled genes isolated from Neurospora crassa are  
 late night- to early morning-specific.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:13096-13101(1996).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RA Shinohara M.L., Bell-Pedersen D., Loros J.J., Dunlap J.C.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF088906; AAC64285.1; -.  
 DR PIR: T47213; T47213.  
 DR GO: GO:0009058; P:biomynthesis; IEA.  
 DR InterPro: IPR001296; Glyco\_transf\_1.  
 DR Pfam: PF00534; Glyco\_transf\_1; 1.  
 DR SEQUENCE 731 AA; 81834 MW; DAF9484FFDF773A2 CRC64;

Query Match 74.0%; Score 37; DB 2; Length 731;  
 Best Local Similarity 75.0%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LMFEDIV 9  
 DB 124 LMLEVDIV 131

RESULT 17  
 08DKV4 PRELIMINARY; PRT; 129 AA.  
 ID 08DKV4  
 AC 08DKV4;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Dihydroneopterin aldolase.  
 GN OrderedlocusNames=tl10747;  
 OS Synecococcus elongatus (Thermosynechococcus elongatus).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.  
 OX NCBI\_TaxID=32046;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BP-1.  
 RX MEDLINE=2225144; PubMed=12240834;  
 RA Nakamura Y., Kaneo T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,  
 RA Watanabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,  
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,  
 RA Shiimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the thermophilic cyanobacterium  
 Thermosynechococcus elongatus BP-1.";  
 RL DNA Res. 9:123-130(2002).  
 DR EMBL: AF005371; BAC08298.1; -.  
 DR HSP; P56740; IDHn.  
 DR GO: GO:0004150; F:dihydroneopterin aldolase activity; IEA.  
 DR GO: GO:0006760; P:folic acid and derivative metabolism; IEA.  
 DR InterPro: IPR006157; FOLB.  
 DR InterPro: IPR006156; FOLB\_fam.  
 DR Pfam: PF02152; FOLB; 1.  
 DR TIGRFAMs: TIGR00525; FOLB; 1.  
 DR TIGRFAMs: TIGR00526; FOLB\_dom; 1.  
 KW Complete proteome.  
 SO SEQUENCE 129 AA; 14567 MW; 689CC2E1C5BEC2D0 CRC64;

Query Match 72.0%; Score 36; DB 2; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WFEIDI 8  
 DB 34 WFEIDI 39

RESULT 18  
 08DBD9 PRELIMINARY; PRT; 264 AA.  
 ID 08DBD9  
 AC 08DBD9;  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Spoom-related protein.  
 GN OrderedlocusNames=V111041;  
 OS Vibrio vulnificus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrrio.  
 OX NCBI\_TaxID=672;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CMCP6;  
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
 RA Choy H.E.;  
 RT "Complete genome sequence of Vibrio vulnificus CMCP6.";  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE016800; AAC09529.1; -.  
 DR InterPro: IPR009776; Spoom.  
 DR Pfam: PF07070; Spoom; 1.  
 KW Complete proteome.  
 SO SEQUENCE 264 AA; 29616 MW; 3CC45B870F2593BC CRC64;

Query Match 72.0%; Score 36; DB 2; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

Qy      2 LMFEID 7
      |||||
Db      213 LMFEID 218

RESULT 19
Q7MGKO PRELIMINARY; PRT; 266 AA.
AC Q7MGKO;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Spoom-related protein.
GN Name=V3231;
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
CC Vibrionaceae; Vibrrio.
OK NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed14656965;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen."
RL Genome Res. 13:2577-2587 (2003).
DR EMBL; AP005343; BAC95995.1; -.
DR InterPro; IPR009776; Spoom.
DR Pfam; PF07070; Spoom; 1.
SQ SEQUENCE 266 AA; 29875 MW; 2C6DC41A32C8F8DC CRC64;

Query Match 72.0%; Score 36; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 LMFEID 7
      |||||
Db      215 LMFEID 220

RESULT 20
OXAI_LACLA STANDARD; PRT; 269 AA.
ID OXAI_LACLA;
AC O9CJ72;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Membrane protein oxa1 precursor.
GN Name=oxa1; OrderedLocustNames=IL0129;
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobactillales; Streptococcaceae; Lactococcus.
OK NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471; DOI=10.1101/gr.169701;
RA Bojotin A., Wincker P., Mauger S., Tallon O., Malarne K.,
RA Weisenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753 (2001).
CC -1- FUNCTION: Required for the insertion of integral membrane proteins
CC into the membrane. May also be involved in protein secretion
CC processes (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the OXA1/OXA family. Subfamily 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006251; AAK04227.1; -.
DR PIR; A86641; A86641.
DR HAMAP; MF_01811; -.
DR InterPro; IPR001708; 60kDa_innermemb.
DR Pfam; PF02096; 60KD IMP; 1.
DR PRINTS; PR00701; 60KDINNERMP.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KM Complete proteome; Lipoprotein; Palmitate; Signal; Transmembrane.
FT SIGNAL 1 20 Potential.
FT CHAIN 21 269 Membrane protein oxa1.
FT TRANSMEM 47 69 Potential.
FT TRANSMEM 166 185 Potential.
FT TRANSMEM 200 222 Potential.
FT LIPID 21 21 N-palmitoyl cysteine (Potential).
FT LIPID 21 21 S-diacylglycerol cysteine (Potential).
SQ SEQUENCE 269 AA; 30294 MW; B57A158CB264F0DF CRC64;

Query Match 72.0%; Score 36; DB 1; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 LMFEID 6
      |||||
Db      155 LMFEID 160

RESULT 21
O87KC8 PRELIMINARY; PRT; 270 AA.
ID O87KC8;
AC O87KC8;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Spoom-related protein.
GN OrderedLocustNames=VP3049;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OK NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RMD 221063 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Nishijima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae."
RL Lancet 361:743-749 (2003).
DR EMBL; AP005083; BAC1312.1; -.
DR InterPro; IPR009776; Spoom.
DR Pfam; PF07070; Spoom; 1.
KW Complete proteome.
SQ SEQUENCE 270 AA; 30310 MW; DFE66C307A0F062D CRC64;

Query Match 72.0%; Score 36; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 LMFEID 7
      |||||
Db      214 LMFEID 219

RESULT 22
O93X27 PRELIMINARY; PRT; 331 AA.
ID O93X27;
AC O93X27;
DT 01-DEC-2001 (TRENBLrel. 19, Created)

```

```

DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)
DE Putative osmosensor histidine kinase (Fragment).
GN Name=unk1;
OS Populus euramericana.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
OC NCBI_TaxID=106131;
OK [1]
RN SEQUENCE FROM N.A.
RP TISSUE=leaves;
RC Caruso A.M.;
RA Theais (2001), Department of Laboratoire de Biologie des Ligneux,
RL Université des Sciences d Orléans, Orléans, France.
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=leaves;
RC Caruso A.M.;
RA Kahlem G., Carpin S.;
RA Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 histidine kinase domain.
DR EMBL; AJ319875; CAC42409.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003559; ATPbind_ATPase.
DR InterPro; IPR004358; Bact_sens_pr_C.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kin_N.
DR InterPro; IPR009082; His_kin_homodim.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; HATPase_1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00389; HisKA_1.
DR PROSITE; PS50109; HIS_KIN_1.
KW Kinase.
FT NON_TER 1 1
FT NON_TER 331 331
SQ SEQUENCE 331 AA; 37258 MW; 2E3CA26ABD5023C4 CRC64;

Query Match 72.0%; Score 36; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LWFEPID 7
DB 181 LWFEPID 186

```

```

RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RA "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli."
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RL FUNCTION: Catalyzes, in vitro, the phosphatidyl group transfer
CC from one phosphatidylglycerol molecule to another to form
CC cardiolipin (CL) (diphosphatidylglycerol) and glycerol (by
CC similarity).
CC -1- CATALYTIC ACTIVITY: 2 Phosphatidylglycerol =
CC diphosphatidylglycerol + glycerol.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (By similarity).
CC -1- SIMILARITY: Belongs to the phospholipase D family. Cardiolipin
CC synthase subfamily.
CC -1- SIMILARITY: Contains 2 PLD phospholipase domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016757; AAN79345.1; -.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR PROSITE; PSS0055; PLD; 2.
KW Complete proteome; Phospholipid biosynthesis; Repeat; Transferase.
FT DOMAIN 108 135 PLD phosphodiesterase 1.
FT DOMAIN 285 312 PLD phosphodiesterase 2.
FT ACT_SITE 113 113 Potential.
FT ACT_SITE 290 290 Potential.
SQ SEQUENCE 413 AA; 47552 MW; 21F9BD2C7CD24CA3 CRC64;

Query Match 72.0%; Score 36; DB 1; Length 413;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FWFEPDI 8
DB 40 FWFEPDI 47

```

```

RESULT 24
YBHO_ECOLI STANDARD; PRT; 413 AA.
ID YBHO_ECOLI
AC P75771;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, last sequence update)
DT 05-JUL-2004 (Rel. 44, last annotation update)
DE Putative cardiolipin synthetase ybho (EC 2.7.8.-) (Cardiolipin
DE synthase) (CL synthase).
GN Name=ybho; OrderedLocNames=b0789, z1008, EG0867;
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562, 83334;
OK [1]
RN SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12."
RT Science 277:1453-1474(1997).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;

```

RA MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Abba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Katihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Samped G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horikuchi T.;  
 RA "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome  
 RT corresponding to the 12.7-28.0 min region on the linkage map.";  
 RL DNA Ref. 3:137-155(1996).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;  
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Knapik H.A.,  
 RA Porfisi G., Hackett J., Klink S., Boutin A., Shao Y., Miller J.,  
 RA Grobeck B.J., Davis N.W., Lim A., Dimalanta E.T., Potamocis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
 RL Nature 409:529-533(2001).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RMD 0509952 / EHEC;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 RA Kunara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
 O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL Nature 409:529-533(2001).  
 [5]  
 RP CHARACTERIZATION.  
 RC STRAIN=HB101;  
 RX MEDLINE=20102623; PubMed=10634942; DOI=10.1016/S1388-1981(99)00193-6;  
 RA Guo D., Tropp B.E.;  
 RT "A second *Escherichia coli* protein with CL synthase activity.";  
 RL Biochim. Biophys. Acta 1483:263-274(2000).  
 CC -1- FUNCTION: Catalyzes, in vitro, the phosphatidyl group transfer  
 from one phosphatidylglycerol molecule to another to form  
 cardiolipin (CL) (diphosphatidylglycerol) and glycerol. Can also  
 catalyze phosphatidyl group transfer to water to form  
 phosphatidate. Catalyzes little, if any, cardiolipin synthesis in  
 vivo even when the expression level is very high.  
 CC -1- CATALYTIC ACTIVITY: 2 Phosphatidylglycerol =  
 diphosphatidylglycerol + glycerol.  
 CC -1- ENZYME REGULATION: Activated by phosphate. Inhibited by  
 cardiolipin and phosphatidate.  
 CC -1- SUBCELLULAR LOCATION: Membrane-bound (Probable).  
 CC -1- SIMILARITY: Belongs to the phospholipase D family. Cardiolipin  
 synthase subfamily.  
 CC -1- SIMILARITY: Contains 2 PLD phosphodiesterase domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB000181; AAC73876.1; -;  
 DR EMBL; D90716; BAA35448.1; -;  
 DR EMBL; AB005259; AAG55160.1; -;  
 DR EMBL; AP002553; BAB34290.1; -;  
 DR PIR; C90737; C90737;  
 DR PIR; D85587; D85587;  
 DR PIR; B64815; B64815;  
 DR ECHOBASE; EB3435; -;  
 DR EcoGene; EG13671; ydhO.  
 DR InterPro; IPR001736; PLD.

DR Pfam; PF00614; PLDC; 2.  
 DR SMART; SM00155; PLDC; 2.  
 DR PROSITE; PS50035; PLD; 2.  
 KW Complete proteome; Phospholipid biosynthesis; Repeat; Transferase.  
 FT DOMAIN 108 135 PLD phosphodiesterase 1.  
 FT DOMAIN 285 312 PLD phosphodiesterase 2.  
 FT ACT SITE 113 113 Potential.  
 FT ACT SITE 290 290 Potential.  
 SQ SEQUENCE 413 AA; 47633 MW; 74998B2A1AD24A11 CRC64;  
 Query Match 72.0%; Score 36; DB 1; Length 413;  
 Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 PLWFEIDI 8  
 Db 40 FIMFEDDV 47  
 RESULT 25  
 Q70XS6 PRELIMINARY; PRT; 445 AA.  
 AC Q70XS6;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE NADH dehydrogenase subunit 4.  
 GN Name=nad4;  
 OS Scutigera coleoptrata (House centipede).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda; Notostrigophora;  
 OC Scutigeroidea; Scutigeridae; Scutigera.  
 OX NCBI\_TaxID=29022;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=14963096;  
 RA Negrisolo E., Minelli A., Valle G.;  
 RT "The mitochondrial genome of the house centipede *Scutigera* and  
 RL myriapod monophyly vs. paraphyly";  
 RL Mol. Biol. Evol. 21:770-780(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=15114420;  
 RA Negrisolo E., Minelli A., Valle G.;  
 RT "Extensive gene order rearrangement in the mitochondrial genome of the  
 centipede *Scutigera coleoptrata*";  
 RL J. Mol. Evol. 58:413-423(2004).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 DR EMBL; AJ507061; CAD45022.2; -;  
 DR GO; GO:0005739; C:mitochondrion; IEA.  
 DR InterPro; IPR003918; NADHdb\_oxred4.  
 DR InterPro; IPR001750; Oxidored\_q1.  
 DR InterPro; IPR00260; Oxidored\_q5\_N.  
 DR Pfam; PF00361; Oxidored\_q1; 1.  
 DR Pfam; PF01059; Oxidored\_q5\_N; 1.  
 DR PRINTS; PR01437; NMOXDRDTSB4.  
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.  
 SQ SEQUENCE 445 AA; 50974 MW; D44F59D0A72D6DB CRC64;  
 Query Match 72.0%; Score 36; DB 2; Length 445;  
 Best Local Similarity 44.4%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 PLWFEIDI 9  
 Db 114 YIMFEISLI 122  
 RESULT 26  
 CAD45022 PRELIMINARY; PRT; 445 AA.  
 ID CAD45022  
 AC CAD45022;  
 DT 10-MAY-2004 (TREMBlrel. 27, Created)



RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buecaw K.H., Scheffer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Datchenko L., Matusik K., Farmer A.A., Rubin G.M., Hong L.,  
RA Strepleon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Frange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [5]  
RN REVIEW ON HPS1 VARIANTS.  
RX MEDLINE=99140254; PubMed=10094567;  
RA Oetting W.S., King R.A.;  
RT "Molecular basis of albinism: mutations and polymorphisms of  
RT pigmentation genes associated with albinism.";  
RL Hum. Mutat. 13:99-115 (1999).  
RN [6]  
RN VARIANT HPS1 ILE-55 DEL.  
RX MEDLINE=98163443; PubMed=9497254;  
RA Oh J., Ho L., Ala-Mello S., Amato D., Armstrong L., Bellucci S.,  
RA Carakushansky G., Ellis J.P., Fong C.-T., Green J.S., Heon E.,  
RA Leguizamo E., Levin A.V., Nieuwenhuis H.K., Plinkers A., Tamura N.,  
RA Whiteford M.L., Yamasaki H., Spritz R.A.;  
RT "Mutation analysis of patients with Hermansky-Pudlak syndrome: a  
RT frameshift hot spot in the HPS gene and apparent locus  
RT heterogeneity.";  
RL Am. J. Hum. Genet. 62:593-598 (1998).  
CC -1- FUNCTION: Component of multiple cytoplasmic organelles. Apparently  
CC crucial for their normal development and function.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Comment=Additional isoforms seem to exist;  
CC Name=I;  
CC IsoId=Q92902-1; Sequence=Displayed;  
CC Name=II;  
CC IsoId=Q92902-2; Sequence=VSP\_004289;  
CC Name=III;  
CC IsoId=Q92902-3; Sequence=VSP\_004290, VSP\_004291;  
CC Name=IV;  
CC IsoId=Q92902-4; Sequence=VSP\_004288;  
CC -1- TISSUE SPECIFICITY: Ubiquitous\_  
CC -1- DISEASE: Defects in HPS1 are the cause of Hermansky-Pudlak  
CC syndrome type 1 (HPS1) [MIM:203300]. HPS1 is an autosomal  
CC recessive disease characterized by oculocutaneous albinism,  
CC bleeding due to platelet storage pool deficiency, and lysosomal  
CC storage defects. This syndrome results from defects of diverse  
CC cytoplasmic organelles including melanosomes, platelet dense  
CC granules and lysosomes. Ceroid storage in the lungs is associated  
CC with pulmonary fibrosis, a common cause of premature death in  
CC individuals with HPS.  
CC -1- DATABASE: NAME=Albinism database (ADB); NOTE=HPS mutations;  
CC WWW="http://www.cbc.umn.edu/rad/".  
CC -1- DATABASE: NAME=Mutations of the HPS gene;  
CC NOTE=Retina International's Scientific Newsletter;  
CC WWW="http://www.retina-international.com/sci-news/hpsmut.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).

CC -----  
DR EMBL; U65576; AAB17869.1; -;  
DR EMBL; U79136; AAB70662.1; -;  
DR EMBL; U79123; AAB70662.1; JOINED.  
DR EMBL; U79124; AAB70662.1; JOINED.  
DR EMBL; U79125; AAB70662.1; JOINED.  
DR EMBL; U79126; AAB70662.1; JOINED.  
DR EMBL; U79127; AAB70662.1; JOINED.  
DR EMBL; U79128; AAB70662.1; JOINED.  
DR EMBL; U79129; AAB70662.1; JOINED.  
DR EMBL; U79130; AAB70662.1; JOINED.  
DR EMBL; U79131; AAB70662.1; JOINED.  
DR EMBL; U79132; AAB70662.1; JOINED.  
DR EMBL; U79133; AAB70662.1; JOINED.  
DR EMBL; U79134; AAB70662.1; JOINED.  
DR EMBL; U79135; AAB70662.1; JOINED.  
DR EMBL; U96721; AAC52074.1; -;  
DR EMBL; BC000175; AAH00175.1; -;  
DR Gene; HGNC:5163; HPS1.  
DR MIM; 604982; -;  
DR MIM; 203300; -;  
DR GO; GO:0005737; C:cytoplasm; TAS.  
DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
DR GO; GO:0005624; C:membrane fraction; TAS.  
DR GO; GO:0005625; C:soluble fraction; TAS.  
DR GO; GO:0007040; P:lysosome organization and biogenesis; TAS.  
DR InterPro; IPR010993; SAM homology.  
KW Albinism; Alternative splicing; Disease mutation;  
KW Hermansky-Pudlak syndrome; Polymorphism; Repeat; Vision.  
KW DOMAIN  
FT 45 650  
FT 1. 4 X 7 AA REPEATS OF [DE]-X-X-X-X-L-L; A  
FT 2. POTENTIAL LYSOSOME/ENDOSOME TRAFFICKING  
FT 3. MOTIF.  
FT 4. Melanosome targeting signal (Potential).  
FT REPEAT 147 153  
FT REPEAT 516 522  
FT REPEAT 644 650  
FT SITE 698 700  
FT VARSPLIC 114 133  
FT VARSPLIC 257 289  
FT VARSPLIC 314 324  
FT VARSPLIC 325 700  
FT VARSPLIC 55 55  
FT VARSPLIC 100 100  
FT VARSPLIC 186 186  
FT VARSPLIC 283 283  
FT VARSPLIC 491 491  
FT VARSPLIC 603 603  
FT VARSPLIC 630 630  
FT VARSPLIC 254 254  
FT CONFLICT 533 537  
FT SEQUENCE 700 AA; 79319 MW; A7872C302B64E0B7 CRC64;  
SQ

Query Match 72.0%; Score 36; DB 1; Length 700;  
Best Local Similarity 75.0%;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
DB 612 FLWRENDW 619

```

RESULT 29
O8WXS5 PRELIMINARY; PRT; 700 AA.
ID O8WXS5;
AC O8WXS5;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)
DE Hermansky-Pudlak syndrome.
Name=HPS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=20256228; PubMed=10798370;
RA Huijing M., Anikster Y., Gahl W.A.;
RT "Characterization of a partial pseudogene homologous to the Hermansky-
Pudlak syndrome gene HPS-1; relevance for mutation detection.";
RL Hum. Genet. 106:370-373(2000).
RN [2]
RP SEQUENCE FROM N.A.
Hermos C.R., Huizing M., Gahl W.A.;
RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF450133; AAL50684.1; -.
DR InterPro: IPR010993; SAM homolog.
SQ SEQUENCE 700 AA; 79291 MW; 0885D3E2D64B3248 CRC64;

Query Match
Best Local Similarity 72.0%; Score 36; DB 2; Length 700;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLWFEIDI 8
Db 612 FLWFENDM 619

RESULT 30
HPS1_MOUSE
ID HPS1_MOUSE STANDARD; PRT; 704 AA.
AC 008983; O35155; O35725; O35950;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, last sequence update)
DT 05-JUL-2004 (Rel. 44, last annotation update)
DE Hermansky-Pudlak syndrome 1 protein homolog.
GN Name=Hps1; Synonyms=Hps, Bp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
STRAIN=BA1B/C; and 129/SvJ;
RC MEDLINE=97301777; PubMed=9158155;
RA Feng G.H., Bailin T., Oh J., Spritz R.A.;
RT "Mouse pale ear (ep) is homologous to human Hermansky-Pudlak syndrome
and contains a rare 'AT-AC' intron.";
RL Hum. Mol. Genet. 6:793-797(1997).
RN [2]
RP SEQUENCE FROM N.A.
STRAIN=C57BL/6; TISSUE=Embryo;
RC MEDLINE=97404379; PubMed=9256466;
RA Gardner J.M., Wildenberg S.C., Keiper N.M., Novak B.K., Rusiniak M.E.,
Swank R.T., Puri N., Fingar J.N., Hagiwara N., Lehman A.L.,
Gales T.U., Bayer M.E., King R.A., Brillant M.H.;
RT "The mouse pale ear (ep) mutation is the homologue of human Hermansky-
Pudlak syndrome.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:9238-9243(1997).
CC -1- FUNCTION: Component of multiple cytoplasmic organelles. Apparently
crucial for their normal development and function. May be involved
in intracellular protein sorting.
CC -1- TISSUE SPECIFICITY: Expressed in all tissues examined with the

```

```

CC possible exception of skeletal muscle. The highest expression was
CC observed in lung, liver, kidney and spleen.
CC -1- DISEASE: Defects in hps1 are the cause of the pale ear (ep) mutant
CC which exhibits hypodysplasia associated with defects of
CC multiple cytoplasmic organelles, including melanosomes, lysosomes,
CC and granular elements of platelets.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U78315; AAB60929.1; -.
DR EMBL; U78966; AAB61333.1; ALT INIT.
DR EMBL; U78955; AAB61333.1; JOINED.
DR EMBL; U78956; AAB61333.1; JOINED.
DR EMBL; U78957; AAB61333.1; JOINED.
DR EMBL; U78958; AAB61333.1; JOINED.
DR EMBL; U78959; AAB61333.1; JOINED.
DR EMBL; U78960; AAB61333.1; JOINED.
DR EMBL; U78961; AAB61333.1; JOINED.
DR EMBL; U78962; AAB61333.1; JOINED.
DR EMBL; U78963; AAB61333.1; JOINED.
DR EMBL; U78964; AAB61333.1; JOINED.
DR EMBL; U78965; AAB61333.1; JOINED.
DR EMBL; U97149; AAB68792.1; -.
DR EMBL; AF003866; AAB68817.1; -.
DR EMBL; AF004352; AAB69159.1; -.
DR EMBL; AF004353; AAB69160.1; -.
DR MGD; MGI:217763; Hps1.
DR GO; GO:0030318; P:melanocyte differentiation; IMP.
DR GO; GO:0060996; P:organelle organization and biogenesis; IMP.
DR InterPro: IPR008946; Strncl_receptor.
KW Aldinism; Polymorphism; Repeat.
FT FT 45
FT FT 51
FT REPEAT 45 51
FT REPEAT 520 526
FT REPEAT 648 654
FT DOMAIN 34 39
FT DOMAIN 252 258
FT DOMAIN 31 31
FT VARIANT 89 89
FT VARIANT 258 258
FT VARIANT 567 567
FT VARIANT 62 62
FT CONFLICT 689 689
FT CONFLICT 689 689
SQ SEQUENCE 704 AA; 79853 MW; CFB8610B3F5A7B89 CRC64;

Query Match
Best Local Similarity 72.0%; Score 36; DB 1; Length 704;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLWFEIDI 8
Db 616 FLWFENDM 623

RESULT 31
O99MK7 PRELIMINARY; PRT; 706 AA.
ID O99MK7;
AC O99MK7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Hermansky-Pudlak syndrome protein.
GN Name=Hps;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21251173; PubMed=11353395;
RA Oh J., Lectrae T.D., Spritz R.A.;
RT "Characterization and evolutionary comparison of rat Hps cDNA and
  exclusion of red-eyed dilution (r) locus.";
RL Mamm. Genome 12:466-468(2001).
DR EMBL; AF333325; AK37597.1; -.
SQ SEQUENCE 706 AA; 80259 MW; 8D325579CBA700E2 CRC64;

Query Match          72.0%; Score 36; DB 2; Length 706;
Best Local Similarity 75.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLWPEIDI 8
DB 618 FLWPEIDM 625

RESULT 32
O935T6 PRELIMINARY; PRT; 76 AA.
AC O935T6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13308;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CNR2368;
RX MEDLINE=95379495; PubMed=7651138;
RA Guedon G., Bourgoin F., Pebay M., Roussel Y., Colmin C., Simonet J.M.,
  Decaris B.;
RT "Characterization and distribution of two insertion sequences, IS1191
  and iso-15981, in Streptococcus thermophilus: does interspecific
  transfer of insertion sequences occur in lactic acid bacteria co-
  cultures?";
RL Mol. Microbiol. 16:69-78(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CNR2368;
RX MEDLINE=20208695; PubMed=10742276;
RA Burrus V., Roussel Y., Decaris B., Guedon G.;
RT "Characterization of a novel integrative element, ICES1, in the
  lactic acid bacterium Streptococcus thermophilus.";
RL Appl. Environ. Microbiol. 66:1749-1753(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CNR2368;
RX MEDLINE=21178499; PubMed=11282600;
RA Burrus V., Bontemps C., Decaris B., Guedon G.;
RT "Characterization of a novel type II restriction-modification system,
  Stb3681, encoded by the integrative element ICES1 of Streptococcus
  thermophilus CNR2368.";
RL Appl. Environ. Microbiol. 67:1522-1528(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CNR2368;
RX MEDLINE=97286550; PubMed=9141697;
RA Roussel Y., Bourgoin F., Guedon G., Pebay M., Decaris B.;
RT "Analysis of the genetic polymorphism between three Streptococcus
  thermophilus strains by comparing their physical and genetic
  organization.";
RL Microbiology 143:1335-1343(1997).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=CNR2368;

```

```

RX MEDLINE=22272761; PubMed=12383726;
RA Burrus V., Pavlovic G., Decaris B., Guedon G.;
RT "The ICES1 element of Streptococcus thermophilus belongs to a large
  family of integrative and conjugative elements that exchange module
  and change their specificity of integration.";
RL Plasmid 48:77-97(2002).
DR EMBL; AJ278471; CAC67543.1; -.
KW Hypothetical protein.
SQ SEQUENCE 76 AA; 8672 MW; 43B935A2F60DFA5F CRC64;

Query Match          70.0%; Score 35; DB 2; Length 76;
Best Local Similarity 55.6%; Pred. No. 69;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWPEIDIV 9
DB 56 FLWPEIDPI 64

RESULT 33
O9WU87 PRELIMINARY; PRT; 112 AA.
AC O9WU87;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Odorant receptor S3 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=99169756; PubMed=10089886;
RA Malnic B., Hirono J., Sato T., Buck L.B.;
RT "Combinatorial receptor codes for odors.";
RL Cell 96:713-723(1999).
DR EMBL; AF121973; AAD27593.1; -.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0004984; F:olfactory receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR Rhodpsn.
DR InterPro; IPR000725; Olfact_receptor.
DR PRINTS; PR00245; OLFACTORYR_
DR PROSITE; PS50262; G_PROTEIN_RECPT_F1_2; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12326 MW; 0FF935D3F7DD71D3 CRC64;

Query Match          70.0%; Score 35; DB 2; Length 112;
Best Local Similarity 55.6%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWPEIDIV 9
DB 26 FLWPEIDPI 34

RESULT 34
O8VEN8 PRELIMINARY; PRT; 194 AA.
AC O8VEN8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Olfactory receptor MOR106-10 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21676683; PubMed-11802173;
RA Zhang X., Firestein S.;
RL "The olfactory receptor gene superfamily of the mouse."
RN Nat. Neurosci. 5:124-133(2002).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE-21664068; PubMed-11875048;
RA Young J.M., Friedman C., Williams E.M., Ross J.A., Tomnes-Priddy L.,
RT Trask B.J.;
RL "Different evolutionary processes shaped the mouse and human olfactory
RT receptor gene families."
RN Hum. Mol. Genet. 11:535-546(2002).
[3]
RP SEQUENCE FROM N.A.
RX Adams M.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY073862; AL61525.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004984; F:olfactory receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR000725; Olfact_receptor.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PRO0245; OLFATORYR.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
KW NON_TER
FT 1
SQ SEQUENCE 194 AA; 22068 MW; 7D96B4D908167C5B CRC64;

Query Match 70.0%; Score 35; DB 2; Length 194;
Best Local Similarity 55.6%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9
DB 40 FLWFPVPII 48

RESULT 35
Q83A27 PRELIMINARY; PRT; 205 AA.
AC Q83A27;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Membrane protein, putative.
GN OrderedLocustNames-CBU1728;
OS Coxiiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OX NCBI_TaxID=777;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=Nine Mile phase I / RSA 493;
MEDLINE=22608657; PubMed=12704232;
RA Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson W.C., Ward N.L., Tettelin J.H., Daviden T.M., Beanan M.J.,
RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA Khouli H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.E., Fraser C.M., Heidelberg J.F.;
RT "Complete genome sequence of the Q-fever pathogen, Coxiella
RT burnetii."
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460(2003).
DR EMBL; AEO16965; AAC01223.1; -
DR TIGR; CRU1728; -
DR GO; GO:0016020; P:membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006419; NMN_transport_PnuC.
DR Pfam; PF04973; NMN_transporter; 1.

```

```

DR TIGRPMs; TIGR01528; NMN_trans_PnuC; 1.
KW Complete proteome.
SQ SEQUENCE 205 AA; 23077 MW; 5304E45265EB4649 CRC64;

Query Match 70.0%; Score 35; DB 2; Length 205;
Best Local Similarity 55.6%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9
DB 144 FLWFIYDAI 152

RESULT 36
Q8IXE0 PRELIMINARY; PRT; 232 AA.
AC Q8IXE0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Seven transmembrane helix receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA Teutsuni S., Aburatani H., Arai K., Akiyama Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB065941; BAC45263.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004984; F:olfactory receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR000725; Olfact_receptor.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PRO0245; OLFATORYR.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor; Transmembrane.
SQ SEQUENCE 232 AA; 26319 MW; 50B0AF3E0B27608F CRC64;

Query Match 70.0%; Score 35; DB 2; Length 232;
Best Local Similarity 77.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLWFEIDIV 9
DB 168 FLWFLIPIV 176

RESULT 37
Y661_ARCFU STANDARD; PRT; 250 AA.
ID Y661_ARCFU
AC Q28612;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical protein AF1661 precursor.
GN OrderedLocustNames=AF1661;
OS Archaeoglobus fulgidus.
OC Archaeae; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
MEDLINE=98049343; PubMed=9389475; DOI=10.1038/37052;
RA Klank H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,

```



RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,  
 RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodex A.,  
 RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,  
 RA Uterback T.R., Cotton M.D., Spriggs T., Arriach P., Kaine B.P.,  
 RA Sykes S.M., Sedow P.W., D'Andrea K.P., Bowman C., Fujii C.,  
 RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,  
 RA Moese C.R., Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 RT reducing archaeon *Archaeoglobus fulgidus*.";  
 RL Nature 390:364-370 (1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AE000989; AAB89610.1; -  
 DR PIR; D69457; D69457.  
 DR TIGR; AF1661; -  
 KM Complete proteome; Hypothetical protein; Signal.  
 FT SIGNAL 1  
 FT CHAIN 26 Potential.  
 SQ SEQUENCE 250 AA; 28240 MW; EBBE5A41C4646C24 CRC64;

Query Match 70.0%; Score 35; DB 1; Length 250;  
 Best Local Similarity 50.0%; Pred. No. 2.3e+02;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLWFEIDI 8  
 Db 101 YLWFAVDV 108

RESULT 38  
 Q8BS40 PRELIMINARY; PRT; 261 AA.  
 AC Q8BS40.  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical conserved protein.  
 GN OrderedLocustNames=O80803;  
 OS Oceanobacillus thelyensis.  
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
 OX NCBI\_TaxID=182710;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HT831;  
 RX MEDLINE=22220767; PubMed=12235376;  
 RA Takami H., Takaki Y., Uchiyama T.;  
 RT "Genome sequence of *Oceanobacillus thelyensis* isolated from the Iheya  
 RT Ridge and its unexpected adaptive capabilities to extreme  
 RT environments.";  
 RL Nucleic Acids Res. 30:3927-3935 (2002).  
 DR EMBL; AP004595; BAC12759.1; -  
 DR GO; GO:0015087; F:covalt ion transporter activity; IEA.  
 DR GO; GO:0006824; P:covalt ion transport; IEA.  
 DR GO; GO:0009236; P:vitamin B12 biosynthesis; IEA.  
 DR InterPro; IPR003339; CblQ.  
 DR Pfam; PF02361; CblQ; 1.  
 KM Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 261 AA; 30165 MW; 145BF98B7DDCA6A CRC64;

Query Match 70.0%; Score 35; DB 2; Length 261;  
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLWFEI 6  
 Db 87 FLWFEV 92

RESULT 39  
 Q9KV00 PRELIMINARY; PRT; 280 AA.  
 AC Q9KV00.  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Spoom-related protein.  
 GN OrderedLocustNames=VC0039;  
 OS Vibrio cholerae.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 CC Vibrionaceae; Vibrrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.B., Read T.D., Tettelin H., Richardson D.L.,  
 RA Ermlaeva M.D., Vamathevan J.U., Base S., Qin H., Dragoi I.,  
 RA Sellers P., McDonald L.A., Uterback T.R., Fleischmann R.D.,  
 RA Niemann W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,  
 RA Mekalanos J.J., Venter J.C., Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*  
 RT cholerae.";  
 RL Nature 406:477-483 (2000).  
 DR EMBL; AB004095; AAF93217.1; -  
 DR PIR; B82372; B82372.  
 DR TIGR; VC0039; -  
 DR InterPro; IPR009776; Spoom.  
 DR Pfam; PF07070; Spoom; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 280 AA; 31302 MW; A3001D4A99C0E2DF CRC64;

Query Match 70.0%; Score 35; DB 2; Length 280;  
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LWFEBID 7  
 Db 229 LWFEBV 234

RESULT 40  
 FPG-COREF STANDARD; PRT; 285 AA.  
 ID FPG-COREF  
 AC O8FP17;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23) (Fapy-DNA  
 DE glycosylase).  
 GN Name=mutM; Synonyms=fpg; OrderedLocustNames=CB1975;  
 OS Corynebacterium efficiens.  
 CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 CC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=152794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
 RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;  
 RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,  
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,  
 RA Gojobori T.;  
 RT "Comparative complete genome sequence analysis of the amino acid  
 RT replacements responsible for the thermostability of *Corynebacterium*  
 RT efficiens.";  
 RL Genome Res. 13:1572-1579 (2003).  
 CC -1- FUNCTION: This enzyme may play a significant role in processes  
 CC leading to recovery from mutagenesis and/or cell death by

```

CC      alkylating agents (By similarity).
CC      -1- CATALYTIC ACTIVITY: Hydrolysis of DNA containing ring-opened N(7)-
CC      methylguanine residues, releasing 2,6-diamino-4-hydroxy-5-(N-
CC      methyl)formamidopyrimidine.
CC      -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC      -1- SIMILARITY: Belongs to the PFG family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.1eb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; AP005220; BAC18785.1; -.
DR      HSSP; P05523; 1K82.
DR      HAMAP; MF_00103; -. 1.
DR      InterPro; IPR000214; Fapy_DNAg_Zn_BS.
DR      InterPro; IPR000191; Fapy_DNA_glyco.
DR      InterPro; IPR010979; Ribosomal_H2TH.
DR      InterPro; IPR010663; ZF-PFG_1IERS.
DR      Pfam; PF01149; Fapy_DNA_glyco; 1.
DR      Pfam; PF06831; H2TH; 1.
DR      Pfam; PF06827; ZF-PFG_1IERS; 1.
DR      ProDom; PD003680; Fapy_DNA_glyco; 1.
DR      TIGRFAMs; TIGR00577; Fpg; 1.
DR      PROSITE; PS01242; PFG; 1.
KM      Complete proteome; DNA repair; Glycosidase; Hydrolase; Zinc;
KW      Zinc-finger.
FT      ZN FING      258      281      Potential.
SQ      SEQUENCE      285 AA; 31162 MW; 600FA9080123B9B1 CRC64;

```

```

Query Match      70.0%; Score 35; DB 1; Length 285;
Best Local Similarity 66.7%; Pred. NO. 2.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 FLWFEIDIV 9
      ||| | | |
Db      62 FLWLELDV 70

```

Search completed: January 12, 2005, 20:14:09  
Job time : 88 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:46:44 ; Search time 21.6 Seconds  
(without alignments)  
27.633 Million cell updates/sec

Title: US-09-870-216C-9  
Perfect score: 45  
Sequence: 1 FLSYDLFVV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	77.8	65	US-08-905-223-453	Sequence 453, App
2	35	77.8	85	US-09-513-999C-5304	Sequence 5304, App
3	35	77.8	162	US-09-663-600A-126	Sequence 126, App
4	35	77.8	162	US-09-663-600A-220	Sequence 220, App
5	35	77.8	163	US-09-149-476-462	Sequence 462, App
6	35	77.8	203	US-09-766-055A-4	Sequence 4, Appli
7	35	77.8	231	US-09-270-767-38981	Sequence 38981, A
8	35	77.8	231	US-09-270-767-54198	Sequence 54198, A
9	35	77.8	382	US-09-248-796A-18423	Sequence 18423, A
10	34	75.6	108	US-09-270-767-45072	Sequence 45072, A
11	34	75.6	195	US-09-270-767-35680	Sequence 35680, A
12	34	75.6	195	US-09-270-767-50897	Sequence 50897, A
13	34	75.6	103	US-09-328-352-5693	Sequence 5693, Ap
14	33	73.3	103	US-09-353-681A-4764	Sequence 4764, Ap
15	33	73.3	370	US-08-837-593-7	Sequence 7, Appli
16	33	73.3	593	US-09-248-796A-19856	Sequence 19856, A
17	31	68.9	132	US-09-270-767-36909	Sequence 36909, A
18	31	68.9	132	US-09-270-767-52126	Sequence 52126, A
19	31	68.9	158	US-09-583-110-3871	Sequence 3871, Ap
20	31	68.9	195	US-09-270-767-38485	Sequence 38485, A
21	31	68.9	195	US-09-270-767-53702	Sequence 53702, A
22	31	68.9	246	US-09-583-110-4242	Sequence 4242, Ap
23	31	68.9	335	US-09-248-796A-17099	Sequence 17099, A
24	31	68.9	346	US-09-107-532A-5546	Sequence 5546, Ap
25	31	68.9	346	US-09-248-796A-15456	Sequence 15456, A
26	31	68.9	349	US-09-328-352-7078	Sequence 7078, Ap
27	31	68.9	400	US-09-107-532A-6266	Sequence 6266, Ap

28	31	68.9	437	4	US-09-540-236-2207	Sequence 2207, Ap
29	31	68.9	451	3	US-09-184-964-3	Sequence 3, Appli
30	31	68.9	453	4	US-09-538-092-472	Sequence 472, App
31	31	68.9	472	4	US-09-486-192-3	Sequence 3, Appli
32	31	68.9	481	3	US-09-537-357-7	Sequence 7, Appli
33	31	68.9	487	3	US-09-537-357-15	Sequence 15, Appli
34	30	66.7	75	4	US-09-134-000C-5626	Sequence 5626, Ap
35	30	66.7	123	4	US-09-248-796A-24081	Sequence 24081, A
36	30	66.7	126	4	US-09-889-463A-6	Sequence 6, Appli
37	30	66.7	130	4	US-09-198-452A-528	Sequence 528, App
38	30	66.7	156	4	US-09-270-767-43152	Sequence 43152, A
39	30	66.7	215	4	US-09-248-796A-28019	Sequence 28019, A
40	30	66.7	324	4	US-09-198-452A-724	Sequence 724, App
41	30	66.7	333	3	US-09-230-637-30	Sequence 30, Appli
42	30	66.7	343	4	US-09-489-039A-10892	Sequence 10892, A
43	30	66.7	346	4	US-09-543-681A-6716	Sequence 6716, Ap
44	30	66.7	349	1	US-08-343-027A-12	Sequence 12, Appli
45	30	66.7	402	4	US-09-107-532A-5454	Sequence 5454, Ap

#### ALIGNMENTS

RESULT 1  
US-08-905-223-453  
Sequence 453, Application US/08905223  
Patent No. 6222029  
GENERAL INFORMATION:  
APPLICANT: Edwards, Jean-Baptiste D.  
APPLICANT: Duclercq, Aymeric  
APPLICANT: Lacroix, Bruno  
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
NUMBER OF SEQUENCES: 503  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/905,223  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-0176  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 453:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 65 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
TISSUE TYPE: Brain  
FEATURE:  
NAME/KEY: s19\_peptide  
LOCATION: -21..-1  
IDENTIFICATION METHOD: Von Heijne matrix  
OTHER INFORMATION: score 5.6  
OTHER INFORMATION: seq LVLSLOPLSLSYD/LF  
US-08-905-223-453

Query Match 77.8%; Score 35; DB 3; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 9.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
 |||||  
 18 LSYDLFV 24

RESULT 2  
 US-09-513-999C-5304  
 ; Sequence 5304, Application US/09513999C  
 ; Patent No. 6783961  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Duclert, A.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 ; Patent No. 6783961  
 ; FILE REFERENCE: 59, US2, REG  
 ; CURRENT APPLICATION NUMBER: US/09/513, 999C  
 ; CURRENT FILING DATE: 2000-02-24  
 ; PRIOR APPLICATION NUMBER: US 60/122,487  
 ; PRIOR FILING DATE: 1999-02-26  
 ; NUMBER OF SEQ ID NOS: 36681  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 5304  
 ; LENGTH: 85  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: 68  
 ; OTHER INFORMATION: Xaa=Ala or Pro or Ser or Thr  
 ; US-09-513-999C-5304

Query Match 77.8%; Score 35; DB 4; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
 |||||  
 18 LSYDLFV 24

RESULT 3  
 US-09-663-600A-126  
 ; Sequence 126, Application US/09663600A  
 ; Patent No. 6573068  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
 ; APPLICANT: Duclert, Aymeric  
 ; APPLICANT: Bougueleret, Lydie  
 ; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS  
 ; FILE REFERENCE: 31, US3, CIP  
 ; CURRENT APPLICATION NUMBER: US/09/663, 600A  
 ; CURRENT FILING DATE: 2000-09-15, 997  
 ; PRIOR APPLICATION NUMBER: 09/191, 997  
 ; PRIOR FILING DATE: 1998-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066, 677  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/069, 957  
 ; PRIOR FILING DATE: 1997-12-17  
 ; PRIOR APPLICATION NUMBER: 60/074, 121  
 ; PRIOR FILING DATE: 1998-02-09  
 ; PRIOR APPLICATION NUMBER: 60/081, 563  
 ; PRIOR FILING DATE: 1998-04-13  
 ; PRIOR APPLICATION NUMBER: 60/096, 116  
 ; PRIOR FILING DATE: 1998-08-10  
 ; PRIOR APPLICATION NUMBER: 60/099, 273  
 ; PRIOR FILING DATE: 1998-09-04  
 ; NUMBER OF SEQ ID NOS: 229  
 ; SOFTWARE: Patent.pm

; SEQ ID NO 126  
 ; LENGTH: 162  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SIGNAL  
 ; LOCATION: -21...-1  
 ; US-09-663-600A-126

Query Match 77.8%; Score 35; DB 4; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
 |||||  
 18 LSYDLFV 24

RESULT 4  
 US-09-663-600A-220  
 ; Sequence 220, Application US/09663600A  
 ; Patent No. 6573068  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
 ; APPLICANT: Duclert, Aymeric  
 ; APPLICANT: Bougueleret, Lydie  
 ; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS  
 ; FILE REFERENCE: 31, US3, CIP  
 ; CURRENT APPLICATION NUMBER: US/09/663, 600A  
 ; CURRENT FILING DATE: 2000-09-15  
 ; PRIOR APPLICATION NUMBER: 09/191, 997  
 ; PRIOR FILING DATE: 1998-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066, 677  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/069, 957  
 ; PRIOR FILING DATE: 1997-12-17  
 ; PRIOR APPLICATION NUMBER: 60/074, 121  
 ; PRIOR FILING DATE: 1998-02-09  
 ; PRIOR APPLICATION NUMBER: 60/081, 563  
 ; PRIOR FILING DATE: 1998-04-13  
 ; PRIOR APPLICATION NUMBER: 60/096, 116  
 ; PRIOR FILING DATE: 1998-08-10  
 ; PRIOR APPLICATION NUMBER: 60/099, 273  
 ; PRIOR FILING DATE: 1998-09-04  
 ; NUMBER OF SEQ ID NOS: 229  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 220  
 ; LENGTH: 162  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SIGNAL  
 ; LOCATION: -94...-1  
 ; US-09-663-600A-220

Query Match 77.8%; Score 35; DB 4; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
 |||||  
 18 LSYDLFV 24

RESULT 5  
 US-09-149-476-462  
 ; Sequence 462, Application US/09149476  
 ; Patent No. 6420526  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: 186 Human Secreted proteins  
 ; FILE REFERENCE: P2002P1  
 ; CURRENT APPLICATION NUMBER: US/09/149, 476

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89</											

```

; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

```

```

Query Match      77.8%; Score 35; DB 4; Length 163;
Best Local Similarity 100.0%; Pred. No. 23;
Matches      7; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

QY      2 LSYDLFV 8
        |||||
DB      18 LSYDLFV 24

```

```

RESULT 6
US-09-766-055A-4
; Sequence 4, Application US/09766055A
; Patent No. 6403354
; GENERAL INFORMATION:
; APPLICANT: XU, SHUANG-YONG
; APPLICANT: SAMUELSON, JAMES
; APPLICANT: PELLETIER, JOHN
; APPLICANT: SIBLEY, MARION
; APPLICANT: WILSON, GEOFFREY G.
; TITLE OF INVENTION: METHOD FOR CLONING AND EXPRESSION OF BcylI RESTRICTION
; TITLE OF INVENTION: ENDOUCLEASE AND BcylI METHYLASE IN E. coli AND
; TITLE OF INVENTION: PURIFICATION OF BcylI AND M.BcylI ENZYMES
; FILE REFERENCE: NEB-185
; CURRENT APPLICATION NUMBER: US/09/766,055A
; CURRENT FILING DATE: 2001-01-19

```

```

; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus Y406
US-09-766-055A-4

```

```

Query Match      77.8%; Score 35; DB 4; Length 203;
Best Local Similarity 75.0%; Pred. No. 29;
Matches      6; Conservative      2; Mismatches      0; Indels      0; Gaps      0;

```

```

QY      1 FLSYDLFV 8
        |||||
DB      136 FLSYDLFV 143

```

```

RESULT 7
US-09-270-767-38981
; Sequence 38981, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38981
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-38981

```

```

Query Match      77.8%; Score 35; DB 4; Length 231;
Best Local Similarity 62.5%; Pred. No. 33;
Matches      5; Conservative      3; Mismatches      0; Indels      0; Gaps      0;

```

```

QY      2 LSYDLFV 9
        |||||
DB      172 LSYDLFV 179

```

```

RESULT 8
US-09-270-767-54198
; Sequence 54198, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54198
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54198

```

```

Query Match      77.8%; Score 35; DB 4; Length 231;
Best Local Similarity 62.5%; Pred. No. 33;
Matches      5; Conservative      3; Mismatches      0; Indels      0; Gaps      0;

```

```

QY      2 LSYDLFV 9
        |||||
DB      172 LSYDLFV 179

```

RESULT 9  
US-09-248-796A-18423  
; Sequence 18423, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 18423  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-18423

Query Match 77.8%; Score 35; DB 4; Length 382;  
Best Local Similarity 66.7%; Pred. No. 55;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFVV 9  
DB 230 FLSYDLSTI 238

RESULT 10  
US-09-270-767-45072  
; Sequence 45072, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 45072  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-45072

Query Match 75.6%; Score 34; DB 4; Length 108;  
Best Local Similarity 66.7%; Pred. No. 24;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLSYDLFVV 9  
DB 86 FFSYDLFVV 94

RESULT 11  
US-09-270-767-35680  
; Sequence 35680, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 35680

; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-35680

Query Match 75.6%; Score 34; DB 4; Length 195;  
Best Local Similarity 44.4%; Pred. No. 43;  
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLFVV 9  
DB 66 YISYELFII 74

RESULT 12  
US-09-270-767-50897  
; Sequence 50897, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 50897  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-50897

Query Match 75.6%; Score 34; DB 4; Length 195;  
Best Local Similarity 44.4%; Pred. No. 43;  
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLFVV 9  
DB 66 YISYELFII 74

RESULT 13  
US-09-328-352-5693  
; Sequence 5693, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 5693  
; LENGTH: 363  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-5693

Query Match 75.6%; Score 34; DB 4; Length 363;  
Best Local Similarity 66.7%; Pred. No. 81;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFVV 9  
DB 39 FLVYDIFMV 47

RESULT 14

```
US-09-543-681A-4764
; Sequence 4764, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4764
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4764

Query Match      73.3%; Score 33; DB 4; Length 103;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

OY      2 LSYDLFV 8
Db      39 ISYDLFV 45

RESULT 15
US-08-837-593-7
; Sequence 7, Application US/08837593
; Patent No. 5977442
; GENERAL INFORMATION:
; APPLICANT: Kleesig, Daniel F.
; APPLICANT: Zhang Zhuqun
; TITLE OF INVENTION: "No. 5977442el Salicylic Acid Induced
; TITLE OF INVENTION: Map Kinase and its Use for Enhanced Disease Resistance in Plant
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dann, Dorfman, Herrell and Skillman,
; ADDRESSER: P.C.
; STREET: 1601 Market Street, Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PasteSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/837,593
; FILING DATE: April 21, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/029,805
; FILING DATE: October 25, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Janet E.
; REGISTRATION NUMBER: 36,252
; REFERENCE/DOCKET NUMBER: Rutgers 97-0016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-248-796A-19856
; Sequence 19856, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19856
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19856

Query Match      73.3%; Score 33; DB 4; Length 593;
Best Local Similarity 77.8%; Pred. No. 2,1e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 FLSYDLFV 9
Db      515 FLSYDLFV 523

RESULT 17
US-09-270-767-36909
; Sequence 36909, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36909
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-36909

Query Match      68.9%; Score 31; DB 4; Length 132;
Best Local Similarity 75.0%; Pred. No. 1,1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 FLSYDLFV 8
Db      82 FLSYDLFV 89
```



RESULT 18  
US-09-270-767-52126  
; Sequence 52126, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1998-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 52126  
; LENGTH: 132  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-52126

Query Match  
Best Local Similarity 68.9%; Score 31; DB 4; Length 132;  
Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 8  
| | | | |  
DB 82 FTSYNLFV 89

RESULT 19  
US-09-583-110-3871  
; Sequence 3871, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 3871  
; LENGTH: 158  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-3871

Query Match  
Best Local Similarity 68.9%; Score 31; DB 4; Length 158;  
Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLF 7  
| | | | |  
DB 63 LSYDLF 68

RESULT 20  
US-09-270-767-38485  
; Sequence 38485, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 38485  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-38485

Query Match  
Best Local Similarity 68.9%; Score 31; DB 4; Length 195;  
Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLSYDLFV 8  
| | | | |  
DB 126 FLKYDSFV 133

RESULT 21  
US-09-270-767-53702  
; Sequence 53702, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1998-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 53702  
; LENGTH: 195  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-53702

Query Match  
Best Local Similarity 68.9%; Score 31; DB 4; Length 195;  
Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLSYDLFV 8  
| | | | |  
DB 126 FLKYDSFV 133

RESULT 22  
US-09-583-110-4242  
; Sequence 4242, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 4242  
; LENGTH: 246  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-4242

Query Match 68.9%; Score 31; DB 4; Length 246;  
Best Local Similarity 77.8%; Pred. No. 2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
||:||||  
Db 133 FLSPFLFV 141

RESULT 23  
US-09-248-796A-17099  
; Sequence 17099, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 17099  
; LENGTH: 335  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-17099

Query Match 68.9%; Score 31; DB 4; Length 335;  
Best Local Similarity 71.4%; Pred. No. 2.8e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLF 7  
||:||||  
Db 270 FLNYDLF 276

RESULT 24  
US-09-107-532A-5546  
; Sequence 5546, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arinello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5546:

SEQUENCE CHARACTERISTICS:  
LENGTH: 344 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium  
FEATURE:

NAME/KEY: misc\_feature  
LOCATION: (B) LOCATION 1...344  
SEQUENCE DESCRIPTION: SEQ ID NO: 5546:  
US-09-107-532A-5546

Query Match 68.9%; Score 31; DB 4; Length 344;  
Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 9  
||:||||  
Db 227 LAHDLFV 234

RESULT 25  
US-09-248-796A-15456  
; Sequence 15456, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 15456  
; LENGTH: 346  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-15456

Query Match 68.9%; Score 31; DB 4; Length 346;  
Best Local Similarity 71.4%; Pred. No. 2.9e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLF 7  
||:||||  
Db 59 FISFDLF 65

RESULT 26  
US-09-328-352-7078  
; Sequence 7078, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7078  
; LENGTH: 349  
; TYPE: PRT

ORGANISM: Acinetobacter baumannii  
US-09-328-352-7078

Query Match 68.9%; Score 31; DB 4; Length 349;  
Best Local Similarity 71.4%; Pred. No. 2.9e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTDLP 7  
DB 279 FLKDYDIF 285

RESULT 27  
US-09-107-532A-6266

Sequence 6266, Application US/09107532A  
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arinello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6266:

SEQUENCE CHARACTERISTICS:

LENGTH: 400 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (B) LOCATION 1..400

SEQUENCE DESCRIPTION: SEQ ID NO: 6266:

US-09-107-532A-6266

Query Match 68.9%; Score 31; DB 4; Length 400;

Best Local Similarity 75.0%; Pred. No. 3.3e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTDLP 8  
DB 197 FLSTGFL 204

RESULT 28

US-09-540-236-2207  
Sequence 2207, Application US/09540236  
Patent No. 6673910

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS

FILE REFERENCE: 2709, 2005-001

CURRENT APPLICATION NUMBER: US/09/540,236

CURRENT FILING DATE: 2000-04-04

NUMBER OF SEQ ID NOS: 3840

SEQ ID NO 2207

LENGTH: 437

TYPE: PRT

ORGANISM: M.catarrhalis

US-09-540-236-2207

Query Match 68.9%; Score 31; DB 4; Length 437;  
Best Local Similarity 75.0%; Pred. No. 3.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLSTDLP 8  
DB 141 FLSTDLP 148

RESULT 29  
US-09-184-964-3

Sequence 3, Application US/09184964

Patent No. 6391574

GENERAL INFORMATION:

APPLICANT: Rine, Jasper D.

APPLICANT: Boyartchuk, Victor L.

TITLE OF INVENTION: APC1 AND RCE1: ISOPRENYLATED CAAX

TITLE OF INVENTION: PROCESSING ENZYMES

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 2200 Sand Hill road, suite 100

CITY: Menlo Park

STATE: CA

COUNTRY: USA

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/184,964

FILING DATE: 03-NOV-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/902,774

FILING DATE: 30-JUL-1997

APPLICATION NUMBER: 60/023,491

FILING DATE: 07-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Suyat, Reginald J.

REGISTRATION NUMBER: 28,172

REFERENCE/DOCKET NUMBER: 09272-006004

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/322-5070

TELEFAX: 650/854-0875

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 451 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: peptide

LOCATION: 1...451

OTHER INFORMATION: /note = "The sequence of the Afcip  
OTHER INFORMATION: protein from yeast presented as a polypeptide sequence"  
us-09-184-964-3

Query Match 68.9%; Score 31; DB 3; Length 451;  
Best Local Similarity 71.4%; Pred. No. 3.7e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTYDLF 7  
DB 87 FLKYDTFTV 93

RESULT 30  
US-09-538-092-472  
Sequence 472, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CuraPatSeqFormatter Version 0.9  
SEQ ID NO 472  
LENGTH: 453  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: Polypeptide Accession Number YCR117W  
US-09-538-092-472

Query Match 68.9%; Score 31; DB 4; Length 453;  
Best Local Similarity 71.4%; Pred. No. 3.8e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTYDLF 7  
DB 87 FLKYDTFTV 93

RESULT 31  
US-09-486-192-3  
Sequence 3, Application US/09486192  
Patent No. 6521440  
GENERAL INFORMATION:  
APPLICANT: Estell, David A.  
TITLE OF INVENTION: Proteases From Gram-Positive Organisms  
FILE REFERENCE: GC386-US  
CURRENT APPLICATION NUMBER: US/09/486,192  
CURRENT FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: PCT/US98/18677  
PRIOR FILING DATE: 1998-09-08  
PRIOR APPLICATION NUMBER: EP9719637.2  
PRIOR FILING DATE: 1997-09-15  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 472  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-486-192-3

Query Match 68.9%; Score 31; DB 4; Length 472;  
Best Local Similarity 71.4%; Pred. No. 3.9e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLSTYDLF 7  
DB 87 FLKYDTFTV 93

RESULT 32  
US-09-537-357-7  
Sequence 7, Application US/09537357  
Patent No. 6271018  
GENERAL INFORMATION:  
APPLICANT: Alan Brash  
TITLE OF INVENTION: MUSKELON (CUCUMIS MELO) HYDROPEROXIDE  
FILE REFERENCE: 06027,0002  
CURRENT APPLICATION NUMBER: US/09/537,357  
CURRENT FILING DATE: 2000-03-29  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 481  
TYPE: PRT  
ORGANISM: Cucumis melo  
US-09-537-357-7

Query Match 68.9%; Score 31; DB 3; Length 481;  
Best Local Similarity 66.7%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FLSTYDLFV 9  
DB 450 FLKYDTFTV 458

RESULT 33  
US-09-537-357-15  
Sequence 15, Application US/09537357  
Patent No. 6271018  
GENERAL INFORMATION:  
APPLICANT: Alan Brash  
TITLE OF INVENTION: MUSKELON (CUCUMIS MELO) HYDROPEROXIDE  
FILE REFERENCE: 06027,0002  
CURRENT APPLICATION NUMBER: US/09/537,357  
CURRENT FILING DATE: 2000-03-29  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 487  
TYPE: PRT  
ORGANISM: Cucumis melo  
NAME/KEY: VARIANT  
LOCATION: (1)...(487)  
OTHER INFORMATION: Xaa = Any Amino Acid  
NAME/KEY: misc\_feature  
LOCATION: (0)...(0)  
OTHER INFORMATION: Accession No. 6271018 AF081955  
US-09-537-357-15

Query Match 68.9%; Score 31; DB 3; Length 487;  
Best Local Similarity 66.7%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FLSTYDLFV 9  
DB 451 FLKYDTFTV 459

RESULT 34  
US-09-134-000C-5626

; Sequence 5626, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5626  
; LENGTH: 75  
; TYPE: PR1  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-5626

Query Match 66.7%; Score 30; DB 4; Length 75;  
Best Local Similarity 44.4%; Pred. No. 93;  
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
| | | | |  
Db 60 FLSFELYIL 68

RESULT 35  
US-09-248-796A-24081  
; Sequence 24081, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196-112  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 24081  
; LENGTH: 123  
; TYPE: PR1  
; ORGANISM: Candida albicans  
US-09-248-796A-24081

Query Match 66.7%; Score 30; DB 4; Length 123;  
Best Local Similarity 62.5%; Pred. No. 1.5e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 8  
| | | | |  
Db 22 FFSYDLYI 29

RESULT 36  
US-09-889-463A-6  
; Sequence 6, Application US/09889463A  
; Patent No. 6680185  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, Saverio C.  
; APPLICANT: Kinney, Anthony J.  
; APPLICANT: Miao, Guo-Hua  
; TITLE OF INVENTION: Plant Polyphenol Oxidase Homologs  
; FILE REFERENCE: BB1330  
; CURRENT APPLICATION NUMBER: US/09/889,463A  
; CURRENT FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/119,590

; PRIOR FILING DATE: 1999-02-10  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 126  
; TYPE: PR1  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (117)  
; OTHER INFORMATION: Xaa = any amino acid  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (120)  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-889-463A-6

Query Match 66.7%; Score 30; DB 4; Length 126;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 8  
| | | | |  
Db 118 FVXFDFV 125

RESULT 37  
US-09-198-452A-528  
; Sequence 528, Application US/09198452A  
; Patent No. 6559294  
; GENERAL INFORMATION:  
; APPLICANT: Griffls, R.  
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, preve  
; FILE REFERENCE: 9710-003-999  
; CURRENT APPLICATION NUMBER: US/09/198,452A  
; CURRENT FILING DATE: 1998-11-24  
; NUMBER OF SEQ ID NOS: 6849  
; SEQ ID NO 528  
; LENGTH: 130  
; TYPE: PR1  
; ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-528

Query Match 66.7%; Score 30; DB 4; Length 130;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 8  
| | | | |  
Db 31 FYSYEIFV 38

RESULT 38  
US-09-270-767-43152  
; Sequence 43152, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1998-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 43152  
; LENGTH: 156  
; TYPE: PR1  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-43152

Search completed: January 12, 2005, 20:17:41  
 Job time : 22.6 secs

Query Match 66.7%; Score 30; DB 4; Length 156;  
 Best Local Similarity 62.5%; Pred. No. 2e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTDLEFV 8  
 : ||: ||  
 Db 135 FVHYDVFV 142

RESULT 39  
 US-09-248-796A-28019  
 ; Sequence 28019, Application US/09248796A  
 ; Patent No. 6747137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keith Weinstock et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.132  
 ; CURRENT APPLICATION NUMBER: US/09/248,796A  
 ; PRIOR FILING DATE: 1999-02-12  
 ; PRIOR APPLICATION NUMBER: US 60/074,725  
 ; PRIOR FILING DATE: 1998-02-13  
 ; PRIOR APPLICATION NUMBER: US 60/096,409  
 ; PRIOR FILING DATE: 1998-08-13  
 ; NUMBER OF SEQ ID NOS: 28208  
 ; SEQ ID NO 28019  
 ; LENGTH: 215  
 ; TYPE: PRT  
 ; ORGANISM: Candida albicans  
 ; FEATURE:  
 ; NAME/KEY: UNSURE  
 ; LOCATION: (21)  
 ; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno  
 US-09-248-796A-28019

Query Match 66.7%; Score 30; DB 4; Length 215;  
 Best Local Similarity 71.4%; Pred. No. 2.7e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSTDLEFV 7  
 : ||: ||  
 Db 79 FVSYELFV 85

RESULT 40  
 US-09-198-452A-724  
 ; Sequence 724, Application US/09198452A  
 ; Patent No. 6559294  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Grifflais, R.  
 ; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
 ; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevent  
 ; TITLE OF INVENTION: and treatment of infection  
 ; FILE REFERENCE: 9710-003-999  
 ; CURRENT APPLICATION NUMBER: US/09/198,452A  
 ; CURRENT FILING DATE: 1998-11-24  
 ; NUMBER OF SEQ ID NOS: 6849  
 ; SEQ ID NO 724  
 ; LENGTH: 324  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia pneumoniae  
 US-09-198-452A-724

Query Match 66.7%; Score 30; DB 4; Length 324;  
 Best Local Similarity 66.7%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTDLEFV 9  
 : ||: ||  
 Db 215 FISHDLAVV 223

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:37:43 ; Search time 201.4 Seconds

(without alignments)  
16.031 Million cell updates/sec

Title: US-09-870-216C-9

Perfect score: 45

Sequence: 1 FLSYDLFPV 9

Scoring table: BLOSUM62

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

1: Geneseqp1980s:\*\n2: Geneseqp1990s:\*\n3: Geneseqp2000s:\*\n4: Geneseqp2001s:\*\n5: Geneseqp2002s:\*\n6: Geneseqp2003as:\*\n7: Geneseqp2003bs:\*\n8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	9	5	ABB08364
2	45	100.0	9	7	ABR82216
3	45	100.0	352	5	ABB08369
4	36	80.0	51	4	AA874753
5	36	80.0	51	5	ABG65310
6	36	80.0	51	8	ADL78577
7	35	77.8	85	3	AAV13169
8	35	77.8	85	2	AAAG01223
9	35	77.8	130	2	AAV17226
10	35	77.8	130	4	AAU38996
11	35	77.8	130	5	ABBS5705
12	35	77.8	160	2	AAV36168
13	35	77.8	162	3	AAV57894
14	35	77.8	162	4	AAV57894
15	35	77.8	162	3	AAV57894
16	35	77.8	162	5	ABR88605
17	35	77.8	162	5	ABR88619
18	35	77.8	162	7	ABR65087
19	35	77.8	162	7	ADD19028
20	35	77.8	162	7	ADJ45973
21	35	77.8	162	8	ADJ46067
22	35	77.8	163	8	ADP19476
23	35	77.8	163	5	AAW74871
24	35	77.8	163	5	ABG95322
25	35	77.8	163	6	ABO34516
				7	AD123177

26	35	77.8	163	8	ADH74179	Adh74179 Human sec
27	35	77.8	177	5	ABB97496	Abb97496 Novel hum
28	35	77.8	203	5	AAE25673	Aae25673 Bacillus
29	35	77.8	203	7	ABR62959	AbR62959 Bacyl res
30	34	75.6	51	4	AAO08931	Aao08931 Human pol
31	34	75.6	55	4	ABP64088	Abp64088 Human ORF
32	34	75.6	98	4	AAV73909	AAV73909 Human col
33	34	75.6	160	7	ADC32867	AdC32867 Human nov
34	34	75.6	247	7	ADM04330	Adm04330 Human pro
35	34	75.6	363	6	ADA34406	Ada34406 Acinetoba
36	34	75.6	412	2	AAW98765	Aaw98765 H. pylori
37	34	75.6	516	5	ABR93691	ABR93691 Herdicia
38	34	75.6	576	7	ADM05144	Adm05144 Human pro
39	34	75.6	674	6	ADA55118	Ada55118 Human pro
40	34	75.6	1031	8	ADM72293	Adm72293 Equine TL
41	33	73.3	90	3	AAV38087	AAV38087 Arabidops
42	33	73.3	103	7	ADP04479	AdP04479 Bacterial
43	33	73.3	119	3	AAQ14935	Aaq14935 Arabidops
44	33	73.3	181	3	AAV38086	AAV38086 Arabidops
45	33	73.3	211	5	ADH32233	Adh32233 Novel yea

## ALIGNMENTS

RESULT 1	
ABB08364	ADH74179
ID ABB08364 standard; protein; 9 AA.	
XX	
AC ABB08364;	
DT 07-MAY-2002 (first entry)	
XX	
DE Synthetic epitope 4 of human cancer antigen eIF3.	
XX	
KW Human; melanoma antigen eukaryotic initiation factor 3; eIF3;	
KW ovarian cancer; MHC; cytosolic; immunomodulator; immune effector cell;	
KW anti-cancer; vaccine.	
XX	
OS Homo sapiens.	
XX	
FH Key	Location/Qualifiers
FT Domain	1 /note= "HLA-2 binding residue"
FT	2 /note= "HLA-2 binding residue"
FT Domain	3..8 /note= "T-cell receptor (TCR) binding domain"
FT	9 /note= "HLA-2 binding residue"
FT Domain	
XX	
WO200192307-A2.	
XX	
PD 06-DEC-2001.	
XX	
PF 30-MAY-2001; 2001WO-US017456.	
XX	
PR 31-MAY-2000; 2000US-0209391P.	
PR 17-AUG-2000; 2000US-0226258P.	
PR 20-DEC-2000; 2000US-0257008P.	
XX	
PA (GENZ ) GENZYME CORP.	
XX	
PI Nicolette CA;	
XX	
DR WPI; 2002-139606/18.	
DR N-PSDB; ABA97215.	
XX	
PT New therapeutic compounds useful against human ovarian cancer, for	
PT modulating immune response in a subject, and for generating antibodies	
PT that specifically recognize and bind to these molecules.	
XX	
PS Claim 29; Page 59; 68bp; English.	





XX New therapeutic compounds useful against human ovarian cancer, for  
PT modulating immune response in a subject, and for generating antibodies  
PT that specifically recognize and bind to these molecules.  
XX  
XX  
PS Claim 9; Page; 68pp; English.

CC The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen E1f3 (melanoma antigen  
 CC leukocyte initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents the human cancer antigen e1f3 variant 4 amino acid sequence.  
 CC Note: This sequence is not present in the specification, but may be  
 CC created from the sequence of the wild-type human cancer antigen e1f3  
 CC sequence given in ABB08360  
 CC  
 CC Sequence 352 AA;  
 CC

Query Match	100.0%	Score 45;	DB 5;	Length 352;
Best Local Similarity	100.0%;	Pred. No. 0.81;		
Matches 9; Conservative	0;	Mismatches	0;	Gaps 0;

Qy	1	FLSYDLFV	9
Db	242	FLSYDLFV	250

RESULT 4  
AAB74753  
ID AAB74753 standard; protein; 51 AA.

AC	AAB74753;
XX	
DT	12-JUN-2001 (first entry)

DE Human secreted protein sequence encoded by gene 21 SEQ ID NO:62.  
XX  
Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;  
KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;  
KW immunostimulant; cytostatic; cardiac; vascular; anti-angiogenic;  
KW ophthalmological; neuroprotectant; nocotropic; anticonvulsant; vaccine;  
KW antialzheimer; antiparkinsonian; antimicrobial; vulvar; gene therapy  
KW immune disorder; hyperproliferative disorder; cardiovascular disease;  
KW cancer; angiogenic disorder; neurological disorder; infectious disease;  
wound healing; regeneration; chemotaxis.

**Homo sapiens.**

PN WO200112775-A2.

PD 22-FEB-2001.

16-AUG-2000; 2000WO-US022325.

PR 17-AUG-1999; 99US-0149182P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI	Rosen CA,	Ni J,	Florence KA,	Fiscella M,	Wei P,	Baker KP;
PI	Birse CE,	Young PE,	Komatsoulis GA,	Moore PA,	Sopnet DR;	

DR WPI; 2001-147550/15.  
DR N-PSDB; AAF81807.  
XX  
Nucleic acid encoding 25 human secreted polypeptides, useful for preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease and diabetic retinopathy.

PS Claim 11; Page 471; 485pp; English

CC AABF178, to AABF1817 encode the human secreted proteins given in AABF4733  
CC to AABF4772. Human secreted proteins can have activities based on the  
CC tissues and cells they are expressed in. Example of activities include:  
CC immunomodulatory; anticancer; dermatological; immunosuppressive;  
CC anti-inflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;  
CC vascular; anti-angiogenic; ophthalmological; neuroprotectant; nootropic;  
CC anticonvulsant; antialzheimers; antiparkinsonian; antimicrobial; and  
CC vaccine. Human secreted proteins can be used in gene therapy and  
CC vaccine. Human secreted protein nucleotide sequences (NM1) and proteins  
CC (PPI) may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate polypeptide expression. For example, NM1  
CC and PPI may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patients genome that  
CC affect the activity of proteins by expressing inactive proteins or to  
CC supplement the patients own production of polypeptides. Disorders that  
CC may be prevented, diagnosed and/or treated include immune disorders,  
CC hyperproliferative disorders (e.g. cancers), cardiovascular diseases,  
CC angiogenic disorders, neurological disorders, infectious diseases and/or  
CC for promoting wound healing, regeneration and /or chemotaxis. AABF178 to  
CC AABF186 and AABF4732 represent sequences used in the exemplification of  
CC the present invention

Query Match	80.0%	Score 36	DB 4	Length 51
Best Local Similarity	66.7%	Pred. No. 7.1		
Matches 6	Conservative 2	Mismatches 1	Indels 0	Gaps 0

```
QY      1 FLSYDLFVV 9
        |::|||::|
Db      7 FVSYDYFIV 15
```

RESULT 5  
ABG65310  
ID ABG65310 standard; protein; 51 AA

DT	27-AUG-2002	(first entry)
XX		
AC	ABG65310;	
zu		

DE Human albumin fusion protein #1965.  
XX  
XX Albumin fusion protein; therapeutic protein X; human albumin; HA  
XX human serum albumin; HSA; cancer; reproductive disorder;  
XX digestive disorder; immune disorder; endocrine disorder;  
XX haemostatic disorder; neural disorder; connective disorder;  
XX cytostatic; antifertility; antiinflammatory; antiulcer;  
XX immunomodulator; anti-HIV; antididiabetic; haemostatic; nootropic;  
XX neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
XX osteopathic; antiarthritic.

**OS Homo sapiens.**

OS Synthetic.

PN WO200177137-

PD 18-OCT-2001.

PF 12-APR-2001; 2001WO-US011988.

PR 12-APR-2000; 2000US-0229358P.

PR 21-DEC-2000; 2000US-0256931P.

Query Match	Best Local	Matches	Similarity	Conservative	Score	DB	Length	Indels	Gaps
QY	1	7	PLSYDLFFV 9	66.7%	36	5	51	1	0
Db	7	7	FVSYDYFIV 15	66.7%	36	5	51	1	0
RESULT 6									
ID	ADL78577		ADL78577 standard; protein; 51 AA.						
AC	ADL78577								
XX	ADL78577								
DT	20-MAY-2004		(first entry)						
DE	Albumin fusion protein related therapeutic protein X, SEQ ID No 2059.								
XX	albumin fusion protein; cytosolic; antihaemic; antiarthritic;								
KW	antialasthmatic; anti-HIV; immunosuppressive; antiinflammatory;								
KW	antipsoiatric; antibacterial; osteopathic; dermatological; antiouc;								
KW	immunomodulator; anticarhythmic; cardiant; nootropic; antilipemic;								
KW	nephrotoxic; utropathic; neuroprotective; antiparkinsonian; tranquilizer;								
KW	antidiabetic; anabolic; hypertensive; vulnerary; gene therapy; cancer;								
KW	reproductive system disorder; therapeutic protein.								
XX	Unidentified.								
OS	US2004010134-A1.								
PN	15-JAN-2004.								
XX	12-APR-2001; 2001US-00833245.								
PF	12-APR-2001; 2001US-00833245.								
XX	12-APR-2000; 2000US-0229358P.								
PR	25-APR-2000; 2000US-019384P.								
XX	21-DEC-2000; 2000US-0256931P.								
PA	(ROSE/) ROSEN C A.								
PA	(HASE/) HASELTINE W A.								
PI	Rosen CA, Haseltine WA;								
DR	WPI; 2004-090519/09.								

PT	New albumin fusion proteins, useful for diagnosing, treating, preventing
PT	or ameliorating diseases or disorders e.g. cancer, anemia, arthritis,
PT	asthma, inflammatory bowel disease or Alzheimer's disease.
XX	
PS	Disclosure; SEQ ID NO 2059; 279pp; English.
XX	
CC	The invention relates to a novel albumin fusion protein. The invention
CC	further relates to: a composition comprising the albumin fusion protein
CC	and a pharmaceutical carrier; a kit comprising the albumin fusion protein
CC	albumin fusion protein formula; a method of treating a disease or
CC	disorder in a patient comprising the step of administering the albumin
CC	fusion protein; a method of treating a patient with a disease or disorder
CC	that is modulated by Therapeutic protein: X, or its fragment or variant;
CC	a method of extending the shelf life of Therapeutic protein: X, or its
CC	fragment or variant; a nucleic acid molecule comprising a polynucleotide
CC	sequence encoding the albumin fusion protein; a vector comprising the
CC	nucleic acid molecule of the albumin fusion protein; and a host cell
CC	comprising the nucleic acid molecule of the albumin fusion protein. The
CC	albumin fusion protein and its compositions have the following
CC	activities: cytostatic, anti-neoplastic, anti-infective, antibacterial,
CC	HIV, immunosuppressive, anti-inflammatory, antiparasitic, antidiabetic,
CC	osteoporotic, dermatological, anti-gout, immunomodulator, antiarrhythmic,
CC	cardiac, neuroprotective, antiparkinsonian, tranquilizer, uropathic,
CC	neuroprotective, and vulnerary. The albumin fusion protein nucleic acid may
CC	be used in gene therapy to treat disorders. The albumin fusion protein is
CC	useful for diagnosing, treating, preventing or ameliorating diseases or
CC	disorders comprising indication: Y. The diseases or disorders include:
CC	cancer (e.g. leukaemia, colon, bone, breast, liver or lung cancer),
CC	immune or haematopoietic diseases (e.g. anaemia, Hodgkin's disease, acute
CC	lymphocytic anaemia, multiple myeloma, arthritis, asthma, AIDS,
CC	autoimmune disease, inflammatory bowel disease, psoriasis or Lyme
CC	disease), reproductive system disorders (e.g. prostatitis, inguinal
CC	hermia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-
CC	leydig tumours), musculoskeletal diseases (e.g. giant cell tumours,
CC	Paget's disease, systemic lupus erythematosus, gout, muscular dystrophy
CC	or cachexia), cardiovascular disease (e.g. rhabdomyoma, heart disease,
CC	arrhythmia, cardiac arrest, heat valve disease, hypernatraemia or
CC	hyponatraemia), mixed foetal diseases (e.g. foetal alcohol syndrome,
CC	Down's syndrome, Patau syndrome, Turner's syndrome, Apert syndrome or Tay-
CC	Sachs disease), excretory diseases (e.g. urinary incontinence, urinary
CC	tract infections or renal disorders), neural or sensory disease (e.g.
CC	Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis,
CC	cerebellar ataxia, attention deficit disorder, autism or obsessive
CC	compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or
CC	occupational lung disease), endocrine diseases (e.g. diabetes, Addison's
CC	disease or glomerulonephritis), digestive diseases (e.g. portal
CC	hypertension, irritable bowel disease, gastric atrophy or pancreatitis)
CC	or connective tissue or epithelial diseases (e.g. Crohn's disease,
CC	scleoderma, wound healing or epidermolysis bullosa). This sequence
CC	represents a therapeutic protein X relating to the albumin fusion protein
CC	of the invention. The sequence listing data for this specification was
CC	downloaded from the USPTO website.
XX	
SQ	Sequence 51 AA;
QY	Query Match 80.0%; Score 36; DB 8; Length 51;
Db	Best Local Similarity 66.7%; Pred. No. 7.1;
	Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
	1 FLSYDLFPV 9
	1:11111111
	7 FVSYDYPTIV 15
RESULT 7	
AAV13169	
ID	AAV13169 standard; protein; 65 AA.
XX	
XX	AAV13169;
DT	22-JUN-1999 (first entry)

```

XX Human secreted protein encoded by 5' EST SEQ ID NO: 163.
DE
XX
XX Human, secreted protein; EST, expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX Homo sapiens.
OS
XX MO906552-A2.
PN
XX 11-FEB-1999.
PD
XX
XX 31-JUL-1998; 98WO-IB001236.
PF
XX
XX 01-AUG-1997; 97US-00905223.
PR
XX (GEST ) GENSET.
PA
XX Dumas Milne Edwards J, Duclert A, Lacroix B;
PI
XX MPI, 1999-153782/13.
DR
XX N-P8DB; AAXS1969.
PT New isolated brain-derived nucleic acids - used to develop products which
PT may have cytokine, immune, regulatory, haematopoiesis regulating, anti-
PT inflammatory or tumour inhibition activity.
PS Claim 34; Page 550; 577pp; English.
XX
XX AAXS1787 to AAXS2019 represent 5' expressed sequence tags (ESTs) for
CC human secreted proteins, and encode the proteins given in AAX12987 to
CC AA113219, respectively. The proteins given represent the signal peptide
CC and an N-terminal fragment of a secreted protein. The nucleic acid
CC sequences can be used for producing secreted human gene products. They
CC can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulation activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used for
CC directing extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell
XX
XX Sequence 65 AA;
SQ
Query Match 77.8%; Score 35; DB 2; Length 65;
Best Local Similarity 100.0%; Pred.No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY 2 LSYDLFV 8
Db 18 LSYDLFV 24
RESULT 8
ID AAG01223 standard; protein; 85 AA.
XX AAG01223;
XX AAG01223;
XX
XX 06-OCT-2000 (first entry)
DT
XX
DE Human secreted protein, SEQ ID NO: 5304.
KW Human, 5' EST, expressed sequence tag; secreted protein; cDNA isolation;

```

XX	gene therapy; chromosome mapping.
XX	
OS	Homo sapiens.
PN	EP1033401-A2.
XX	
PD	06-SEP-2000.
XX	
PF	21-FEB-2000; 2000EP-00200610.
XX	
PR	26-FEB-1999; 99US-0122487P.
XX	
PA	(GEST ) GENSET.
PI	Dumas Milne Edwards J, Duclert A, Giordano J,
DR	WPI; 2000-500381/45.
DR	N-PSDB; AAC01229.
XX	
PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT	diagnostic, forensic, gene therapy and chromosome mapping procedures.
PS	Claim 13; SEQ ID NO 5304; 71pp + Sequence Listing; English.
XX	
CC	The present sequence is a polypeptide encoded by one of a large number of
CC	5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
CC	prepared from total human RNAs or polyA+ RNAs derived from 30 different
CC	tissues. EST sequences usually correspond mainly to the 3' untranslated
CC	region (UTR) of the mRNA because they are often obtained from oligo-dT
CC	primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
CC	sequences derived from the 5' ends of mRNAs and even in those cases where
CC	longer cDNA sequences have been obtained, the full 5' UTR is rarely
CC	included. 5' ESTs are derived from mRNAs with intact 5' ends and can
CC	therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
CC	are also used in diagnostic, forensic, gene therapy and chromosome
CC	mapping procedures. They are used to obtain upstream regulatory sequences
CC	and to design expression and secretion vectors
XX	
SQ	Sequence 85 AA;
XX	
Query Match	77.8%; Score 35; DB 3; Length 85;
Best Local Similarity	100.0%; Pred. No. 20;
Matches 7; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	2 LSYDLPV 8
DB	18 LSYDLPV 24
XX	
RESULT 9	
AA17226	
ID	AA17226 standard; protein: 130 AA.
XX	
AC	AA17226;
XX	
DT	09-AUG-1999 (first entry)
XX	
DE	Human secreted protein (clone pe204-1).
XX	
KW	Secreted protein; kidney; lung; brain; blood; testis; bone marrow;
KW	nutritional activity; cytokine; cell proliferation; immune stimulation;
KW	hematopoiesis regulation; tissue growth; thrombolytic; gene therapy;
KW	anti-inflammatory; tumour invasion.
XX	
OS	Homo sapiens.
XX	
PN	W09926961-A1.
XX	
PD	03-JUN-1999.
XX	
PF	24-NOV-1998; 98WO-US025149.
XX	

PR	26-NOV-1997;	97US-0066804P.
PR	23-NOV-1998;	98US-00197886.
XX		
PA	(GENMY ) GENETICS INST INC.	
PI	Jacobs K, McCoy JM, Lavallic ER, Collins-Racie LA, Evans C,	
PI	Meeberg D, Treacy M, Agostino MJ, Steinhinger RJ, Wong GG, Clark HF,	
PI	Fechtel K;	
DR	WPI; 1999-357809/30.	
DR	N-PSDB; AAX60808.	
XX		
PT	New polynucleotides encoding secreted proteins.	
XX		
PS	Claim 34; Page 123; 133pp; English.	
XX		
CC	The invention relates to secreted proteins (AAV17219-228) encoded by	
CC	polynucleotides obtained from human fetal kidney, adult lung, adult	
CC	kidney, adult brain, adult blood, adult testes, and fetal brain and	
CC	murine adult bone marrow cDNA libraries. The secreted protein nucleic acid	
CC	sequences (X6801-811) correspond to clones bd306-7, gJ283-6, FK317-3,	
CC	k213-2x, na31c-1, nE33-20, np204-1, pE204-1, Yal-1 and yB-1, (all clones	
CC	are deposited as ATCC 95599). The PNs and proteins are predicted to have	
CC	biological activities which would make them suitable for treating,	
CC	preventing or ameliorating medical conditions in humans and animals,	
CC	although no supporting data is given. Suggested activities include	
CC	nutritional activity, cytokine and cell proliferation/differentiation	
CC	activity, immune stimulating (e.g. as vaccines) or suppressing activity,	
CC	hematopoiesis regulating activity, tissue growth activity, activin/ inhibin activity, chemocactive/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, and tumour inhibition activity. The PNs are also stated to be useful for gene therapy	
CC		
XX		
SQ	Sequence 130 AA;	
Query Match	77.8%; Score 35; DB 2; Length 130;	
Best Local Similarity	100.0%; Pred. No. 31;	
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	2 LSVDLFV 8	
Db	18 LSYDLFV 24	
RESULT 10		
AU38996	AAU38996 standard; protein; 130 AA.	
AC	AAU38996;	
DT	16-JAN-2002 (first entry)	
DE	Human secreted protein pe241_1.	
XX		
KM	Human; secreted protein; antiinflammatory; immunosuppressive; neurotropic; neuroprotective; antiarthritic; antimicrobial; vulnerary; cytoskeletal; antidiabetic; virucide; antifertility; anticoagulant; vasotropic; antiasthmatic; immunostimulant; dermatological; antihematitic; antitumor; antiviral; osteopathic; tranquiliser; cerebroprotective; cytokine; cell proliferation; cell differentiation; immune deficiency; severe combined immunodeficiency; SCID; tumour; autoimmune disorder; multiple sclerosis; rheumatoid arthritis; graft-versus-host disease; myeloid deficiency; wound healing; ulcer; periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease; Parkinson's disease; Huntington's disease; infection; cardiac disease; stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen; food supplement; vaccine.	
OS	Homo sapiens.	
XX		
PN	WO200175068-A2.	

PD 11-OCT-2001.  
 PF 22-MAR-2001; 2001WO-US009369.  
 XX  
 PR 30-MAR-2000; 2000US-00539330.  
 XX 04-DEC-2000; 2000US-00729674.  
 PA (GENY ) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoy JM, Lavallie E, Collins-Racie LA, Evans C;  
 PI Treacy M, Agostino MJ, Seelinger RJ, Spaulding V, Wong GG, Clark H;  
 PI Fechtel K, Nerberg D;  
 XX WPI; 2001-639363/73.  
 DR N-PSDB; AAS59214.  
 XX  
 PT Secreted human proteins, useful as vaccine for treating various diseases  
 PT such as autoimmune disorders (e.g. multiple sclerosis), and nervous  
 PT system disorders (e.g. stroke).  
 PS  
 PS Disclosure; Page 468-469; 619pp; English.  
 XX  
 CC The invention relates to novel human secreted proteins, the nucleic acids  
 CC encoding them. The protein may exhibit cytokine, cell proliferation or  
 CC cell differentiation activity or may induce production of other cytokines  
 CC in certain cell populations and may exhibit immune stimulating or immune  
 CC suppressing activity, which is useful for the treatment of various immune  
 CC deficiencies and disorders e.g. severe combined immunodeficiency (SCID),  
 CC autoimmune disorders e.g. multiple sclerosis, systemic lupus  
 CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation.  
 CC The proteins are also useful in the treatment of diseases and disorders  
 CC including tissue, skin and organ transplantation and in graft-versus-host  
 CC diseases (GVHD), in the induction of tumour immunity, myeloid or lymphoid  
 CC cell deficiencies, wound healing and tissue repair, in the treatment of  
 CC burns, incisions and ulcers; as well as in treatment of periodontal  
 CC disease, osteoporosis or osteoarthritis, mediated by inflammatory  
 CC processes, diseases of the peripheral nervous system, Alzheimer's,  
 CC Parkinson's disease, Huntington's disease, amyotrophic lateral  
 CC sclerosis, and Shy-Drager syndrome, infections, infarction of cardiac and  
 CC central nervous system vessel e.g. stroke, sepsis, inflammatory bowel  
 CC disease, ulcers, bone regeneration. The protein, having activin- or  
 CC inhibin-related activities is useful as a contraceptive based on the  
 CC ability of inhibiting to decrease fertility in female mammals and decrease  
 CC spermatogenesis in male mammals. The proteins and nucleic acids are also  
 CC useful as food supplements. The present sequence represents a secreted  
 CC protein of the invention  
 CC  
 XX  
 XX Sequence 130 AA;  
 SQ  
 Query Match 77.8%; Score 35; DB 4; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LSYDLFV 8  
 |||||  
 18 LSYDLFV 24  
 DB  
 RESULT 11  
 ABB55705  
 ID ABB55705 standard; protein; 130 AA.  
 XX  
 XX ABB55705;  
 AC  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 16.  
 XX  
 KW Human; clone h3306-7; clone yb8-1; ATCC number 98599; gene therapy;  
 KW immune disorder; bacterial infection; fungal infection; cancer; tumour;  
 KW autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;  
 KW osteoporosis; osteoarthritis; nervous system disorder; neuropathy;

KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;  
 KW haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;  
 KW ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;  
 KW Crohn's disease; cytosolic; anti-inflammatory; immunomodulator;  
 KW neutropenic; haemostatic; thrombolytic; anti-inflammatory.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2001039335-A1.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PE 04-DEC-2000; 2000US-00729674.  
 XX  
 PR 26-NOV-1997; 97US-0126425P.  
 PR 04-DEC-1997; 97US-0067454P.  
 PR 20-DEC-1997; 97US-0068379P.  
 PR 02-JAN-1998; 98US-0070346P.  
 PR 07-JAN-1998; 98US-0070643P.  
 PR 08-JAN-1998; 98US-0070755P.  
 PR 13-JAN-1998; 98US-0071304P.  
 PR 22-JAN-1998; 98US-0072134P.  
 PR 30-JAN-1998; 98US-0073095P.  
 PR 18-FEB-1998; 98US-0075038P.  
 PR 23-NOV-1998; 98US-00197886.  
 PR 30-MAR-2000; 2000US-00539330.  
 XX  
 PA (JACO/) JACOBS K.  
 PA (MCCO/) MCCOY J M.  
 PA (LAVAN/) LAVALLIE E R.  
 PA (COLL/) COLLINS-RACIE L A.  
 PA (EVAN/) EVANS C.  
 PA (MERB/) MERBERG D.  
 PA (TREAC/) TREACY M.  
 PA (AGOS/) AGOSTINO M J.  
 PA (STREI/) STEININGER R J.  
 PA (SPAU/) SPAUDING V.  
 PA (WONG/) WONG G G.  
 PA (CLAR/) CLARK H.  
 PA (FECH/) FECHTEL K.  
 XX  
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C,  
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V,  
 PI Wong GG, Clark H, Fechtel K;  
 XX  
 DR MPI: 2002-040725/05.  
 DR N-PSDB; ABA90883.  
 XX  
 PT New secreted proteins and encoding polynucleotides, useful in gene  
 PT therapies, particularly for preventing or treating autoimmune disorders,  
 PT cancer, graft-versus-host disease, wound, osteoporosis, stroke or  
 PT inflammations.  
 XX  
 PS Disclosure; Page 189-190; 349pp; English.  
 XX  
 CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and  
 CC ABA90980) and encoded proteins (ABBS5698-ABBS5800), especially  
 CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and  
 CC proteins SEQ ID NO 2 (ABBS5698) and SEQ ID NO 20 (ABBS5707) contained in  
 CC clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1  
 CC are deposited with the American Type Culture Collection (ATCC) with  
 CC accession number 98599. The polynucleotides and encoded polypeptides have  
 CC cytosolic, anti-inflammatory, immunomodulator, vulnery,  
 CC neutropenic, anti-inflammatory, immunomodulator, vulnery,  
 CC and anti-inflammatory activity and acting as cytokine modulators,  
 CC haematopoiesis regulators, tissue growth modulators and/or cadherin  
 CC suppressors. The polypeptides and polynucleotides are useful in gene  
 CC therapies, particularly for preventing, treating or ameliorating any of  
 CC the following diseases: immune deficiency disorders; e.g. bacterial  
 CC or fungal infections, autoimmune disorders, cancer, systemic lupus  
 CC erythematosus or graft-versus-host disease, myeloid or lymphoid cell  
 CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or  
 CC osteoarthritis; central and peripheral nervous system diseases and

CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's  
 CC disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;  
 CC haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis  
 CC or systemic inflammatory response syndrome, ischaemia-reperfusion injury,  
 CC endotoxin lethality, arthritis, inflammatory bowel disease or Crohn's  
 CC disease; or tumours or cancers, pemphigus vulgaris or pemphigus foliaceus  
 XX  
 SQ Sequence 130 AA;  
 XX  
 OY Query Match 77.8%; Score 35; DB 5; Length 130;  
 DB Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 LSVDLFV 8  
 DB 18 LSVDLFV 24  
 XX  
 RESULT 12  
 AAY36168  
 ID AAY36168 standard; protein; 160 AA.  
 XX  
 AC AAY36168;  
 XX  
 DT 23-SEP-1999 (first entry)  
 XX  
 DE Human secreted protein.#40.  
 XX  
 KW Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;  
 KW diagnostic; gene therapy; chromosome mapping; secretion vector.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO925825-A2.  
 XX  
 PD 27-MAY-1999.  
 XX  
 PF 13-NOV-1998; 98WO-IB001862.  
 XX  
 PR 13-NOV-1997; 97US-0066677P.  
 PR 17-DEC-1997; 97US-0069957P.  
 PR 09-FEB-1998; 98US-0074121P.  
 PR 13-APR-1998; 98US-0081563P.  
 PR 10-AUG-1998; 98US-0096116P.  
 PR 04-SEP-1998; 98US-0099273P.  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Bougueleret L, Duclert A, Dumas Milne Edwards J;  
 XX  
 DR MPI: 1999-347472/29.  
 DR N-PSDB; AAX97852.  
 XX  
 PT Extended cDNAs encoding secreted proteins.  
 XX  
 PS Example 28; Page 247-248; 307pp; English.  
 XX  
 CC AA36129-Y36222 represent novel human secreted proteins encoded by the  
 CC extended cDNA sequences represented in AAX97813-X97906. The proteins of  
 CC the invention have cytosolic, thrombotic and osteopathic activity. The  
 CC extended cDNAs can be used to express secreted proteins or parts of them  
 CC or to obtain antibodies capable of binding to the secreted proteins. They  
 CC may also be used in diagnostic, forensic, gene therapy and chromosome  
 CC mapping procedures. Uses also include design of expression vectors and  
 CC secretion vectors  
 XX  
 SQ Sequence 160 AA;  
 XX  
 OY Query Match 77.8%; Score 35; DB 2; Length 160;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 LSVDLFV 8

Db 18 LSYDLFV 24

# RESULT 13

AA536215 AAY36215 standard; protein; 162 AA.

AC AAY36215;

DT 23-SEP-1999 (first entry)

DE Human secreted protein #87.

XX Secreted protein; human; cytosolic; thrombotic; osteopathic; forensic;

KW diagnostic; gene therapy; chromosome mapping; secretion vector.

OS Homo sapiens.

PN WO925825-A2.

PD 27-MAY-1999.

PF 13-NOV-1998; 98WO-IB001862.

PR 13-NOV-1997; 97US-0066677P.

PR 17-DEC-1997; 97US-0069577P.

PR 09-FEB-1998; 98US-0074121P.

PR 13-APR-1998; 98US-0081563P.

PR 10-AUG-1998; 98US-0096116P.

PR 04-SEP-1998; 98US-0092273P.

XX (BEST) GENSET.

XX Bougueleret L, Duclert A, Dumas Milne Edwards J;

XX MPI; 1999-347472/29.

XX N-PSDB; AAY37899.

XX Extended cDNAs encoding secreted proteins.

XX Claim 7; Page 304-305; 307pp; English.

XX AAY36129-Y36222 represent novel human secreted proteins encoded by the

CC extended cDNA sequences represented in AAY37813-A37906. The proteins of

CC the invention have cytosolic, thrombotic and osteopathic activity. The

CC extended cDNAs can be used to express secreted proteins or parts of them

CC or to obtain antibodies capable of binding to the secreted proteins. They

CC may also be used in diagnostic, forensic, gene therapy and chromosome

CC mapping procedures. Uses also include design of expression vectors and

CC secretion vectors

XX Sequence 162 AA;

XX SQ

XX Query Match 77.8%; Score 35; DB 2; Length 162;

XX Best Local Similarity 100.0%; Pred. No. 39;

XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Db 2 LSYDLFV 8

18 LSYDLFV 24

# RESULT 14

AA57894 AAY57894 standard; protein; 162 AA.

AC AAY57894;

DT 23-MAR-2000 (first entry)

DE Human transmembrane protein HTPN-18.

KW Human; transmembrane protein; HTPN; diagnosis; immunospecific;

XX antiproliferative; neuroprotective; immune disorder;

XX reproductive disorder; smooth muscle disorder; neurological disorder;

XX Gastrointestinal disorder; developmental disorder;

XX cell proliferative disorder.

XX Homo sapiens.

OS MO961471-A2.

PN 02-DEC-1999.

PD 28-MAY-1999; 99WO-US011904.

PF 29-MAY-1998; 98US-0087260P.

PR 02-JUL-1998; 98US-0091674P.

PR 02-OCT-1998; 98US-0102854P.

PR 24-NOV-1998; 98US-0109869P.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;

XX Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR, Au-Young J;

XX MPI; 2000-072605/06.

XX N-PSDB; AAY56715.

XX Proteins, polynucleotides, vectors, host cells and antibodies used to

XX diagnose, treat or prevent immune, reproductive, smooth muscle,

XX PT neurological, gastrointestinal, developmental and cell proliferative

XX disorders.

XX Claim 1; Page 125; 229pp; English.

XX AAY56698 to AAY56776 encode AAY57877 to AAY57955 which represent human

CC transmembrane proteins designated HTPN-1 to HTPN-79, respectively. The

CC transmembrane protein have immunospecific, antiproliferative and

CC neuroprotective activities. The human transmembrane proteins,

CC polynucleotides encoding them and other compositions and methods from the

CC present invention, can be used for the diagnosis, treatment or prevention

CC of immune, reproductive, smooth muscle, neurological, gastrointestinal,

CC developmental and cell proliferative disorders. The HTPN's can be used

CC to treat or prevent disorders associated with a decreased expression or

CC activity of HTPN

XX Sequence 162 AA;

XX SQ

XX Query Match 77.8%; Score 35; DB 3; Length 162;

XX Best Local Similarity 100.0%; Pred. No. 39;

XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Db 2 LSYDLFV 8

18 LSYDLFV 24

# RESULT 15

AA88605 AAB88605 standard; protein; 162 AA.

AC AAB88605;

DT 04-JUN-2001 (first entry)

DE Human hydrophobic domain containing protein clone HP10773 #129.

XX Human; hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;

XX antianemic; vulnary; antiulcer; osteopathic; anti-inflammatory;

XX cytosolic; gene therapy; autoimmune disorder; multiple sclerosis;

XX HIV infection; anaemia; burn; ulcer; osteoporosis; tumour; wound healing;

XX inflammatory bowel disease; nutritional supplement; appetite; vaccine;

XX behavioural characteristic; immune response.

OS Homo sapiens.  
 XX  
 XX WO200112660-A2.  
 XX  
 XX 22-FEB-2001.  
 XX  
 XX 10-AUG-2000; 2000WO-JP005356.  
 XX  
 XX 17-AUG-1999; 99JP-00230344.  
 PR 07-SEP-1999; 99JP-00252551.  
 PR 01-OCT-1999; 99JP-00281132.  
 PR 22-OCT-1999; 99JP-00301624.  
 PR 04-NOV-1999; 99JP-00313877.  
 XX  
 XX (SAGA ) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX  
 PI Kato S, Kimura T;  
 XX  
 DR WPI; 2001-160059/16.  
 DR N-PSDB; AAF94505.  
 XX  
 PT Human proteins with hydrophobic domains and the DNAs which encode them  
 PT are useful for treating autoimmune disorders, burns and tumors and for  
 PT screening novel pharmaceuticals.  
 PT  
 PT  
 XX  
 XX Claim 1; Page 478-479; 518pp; English.  
 PS  
 XX AAF94417 to AAF94516 encode the human proteins given in AAB88557 to  
 CC AAB88606 (I) which have a hydrophobic domain. (I) have immunosuppressant,  
 CC anti-HIV, neuroprotective, antianemic, vulnerary, antitumor,  
 CC osteopathic, anti-inflammatory and cytostatic activities, and can be used  
 CC in gene therapy. (II) can be used as pharmaceuticals and as antigens to  
 CC prepare antibodies. DNA and cDNA (II) encoding (I) can be used as probes  
 CC for genetic diagnosis and gene sources for gene therapy or for producing  
 CC (I) in large quantities. Cells containing (II) are used for the detection  
 CC of ligands or receptors corresponding to membrane or secretory proteins  
 CC and to screen small molecule novel pharmaceuticals. Antibodies directed  
 CC to (I) can be used for the detection, quantification and purification of  
 CC (I). Activities of (I) may include cytokine and cell  
 CC proliferation/differentiation function, immune stimulating or suppressing  
 CC activity, haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, receptor/ligand activity and anti-inflammatory  
 CC activity. (I) and (II) can be used to treat autoimmune disorders e.g.  
 CC multiple sclerosis, HIV infections, anaemia, burns, ulcers, osteoporosis,  
 CC inflammatory bowel disease and tumours. (I) and (II) can also be used for  
 CC wound healing, as nutritional sources or supplements e.g. as amino acid,  
 CC carbon or nitrogen source, to effect metabolism, catabolism, anabolism,  
 CC processing and utilisation of dietary fat, protein, carbohydrate,  
 CC vitamins and minerals, to effect behavioural characteristics, to affect  
 CC appetite, and can act as antigens in vaccines to raise an immune response  
 CC to the protein or another material cross-reactive with the protein  
 CC  
 XX  
 XX Sequence 162 AA;  
 SQ  
 QY Query Match 77.8%; Score 35; DB 4; Length 162;  
 Db Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 LSYDLFV 8  
 18 LSYDLFV 24  
 RESULT 16  
 ABB89619  
 ID ABB89619 standard; protein; 162 AA.  
 XX  
 XX ABB89619;  
 AC  
 XX 24-MAY-2002 (first entry)  
 DT  
 XX

DE Human polypeptide SEQ ID NO 1995.  
 XX  
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antifungal;  
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200190304-A2.  
 XX  
 XX 29-NOV-2001.  
 XX  
 XX 18-MAY-2001; 2001WO-US016450.  
 XX  
 XX 19-MAY-2000; 2000US-0205515P.  
 XX  
 XX (HDMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX Birse CE, Rosen CA;  
 XX  
 XX WPI; 2002-122018/16.  
 DR N-PSDB; ABL90028.  
 XX  
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
 PT prevention of neural, immune system, muscular, reproductive,  
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
 PT disorders.  
 PT  
 XX  
 XX Claim 11; SEQ ID NO 1995; 2081pp + Sequence Listing; English.  
 PS  
 XX The invention relates to novel genes (ABL89449-ABL90853) and proteins  
 CC (AAB89040-ABB90444) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infectious diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 162 AA;  
 SQ  
 QY Query Match 77.8%; Score 35; DB 5; Length 162;  
 Db Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 LSYDLFV 8  
 18 LSYDLFV 24  
 RESULT 17  
 ABB65087  
 ID ABB65087 standard; protein; 162 AA.  
 XX  
 XX ABB65087;  
 AC  
 XX 12-NOV-2002 (first entry)  
 DT  
 XX Hypoxia-repressed protein #10.  
 XX  
 XX Cytostatic; vasotropic; tranquiliser; antithrombotic; gene therapy;  
 KW antiinflammatory; vulnerary; gynecological; ophthalmological; vaccine;  
 KW

KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;  
KM ischaemic condition; reperfusion injury; retinopathy; neonatal stress;  
KM pre-eclampsia; atherosclerosis; inflammatory condition; wound healing;  
KM inflammation; erythropoiesis; hair loss; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200246465-A2.  
XX  
PD 13-JUN-2002.  
XX  
PF 10-DEC-2001; 2001WO-GB005458.  
XX  
PR 08-DEC-2000; 2000GB-00030076.  
PR 08-FEB-2001; 2001GB-00003156.  
PR 25-OCT-2001; 2001GB-00025666.  
XX  
PA (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX  
PI White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;  
PI Rayner WN;  
XX  
DR WPI; 2002-627238/67.  
XX  
PT Identifying a gene involved in disease for treating hypoxia-regulated  
PT conditions, comprises comparing the transcriptome/proteome of two cell  
PT types under different conditions and identifying a differentially  
PT regulated gene.  
XX  
PS Claim 13; Page 289; 538pp; English.  
XX  
CC The present invention relates to methods for identifying genes and  
CC proteins that are implicated in a specific disease or physiological  
CC condition. The method comprises comparing the transcriptome/proteome of a  
CC specialised cell type implicated in a disease or condition with that of a  
CC second specialised cell type, under two experimental conditions, and  
CC identifying a gene that is differentially regulated in the two  
CC specialised cell types under experimental conditions. ABV77873-ABV78116  
CC and ABP5061-ABP5257 were identified using the methods of the invention.  
CC The coding sequences and proteins are useful for treating a disease in a  
CC patient, for manufacture of a medicament for treating hypoxia-regulated  
CC conditions, and for regulating tumorigenesis, angiogenesis, apoptosis,  
CC biological response to hypoxia conditions, or hypoxic-associated  
CC pathology in a patient. The coding sequences and proteins are also useful  
CC for monitoring the therapeutic treatment of a disease or physiological  
CC condition, such as cancer, ischaemic conditions, reperfusion injury,  
CC retinopathy, neonatal stress, pre-eclampsia, atherosclerosis, inflammatory  
CC conditions, wound healing, inflammation, erythropoiesis or hair loss  
CC  
XX  
SQ Sequence 162 AA;  
XX  
Query Match 77.8%; Score 35; DB 5; Length 162;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 LSYDLFV 8  
DB 18 LSYDLFV 24  
XX  
RESULT 18  
ADD19028  
ID ADD19028 standard; protein; 162 AA.  
XX  
AC ADD19028;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Human disease related protein SegID517.  
XX  
KW human; disease state; cytostatic; anti-inflammatory; ophthalmological;  
KW antiarteriosclerotic; vulnerary; gene therapy;  
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;

KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;  
KW glucose transportation; catecholamine synthesis; iron transport;  
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;  
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;  
KW inflammatory condition; wound healing.  
XX  
OS Homo sapiens.  
XX  
PN WO2003018621-A2.  
XX  
PD 06-MAR-2003.  
XX  
PF 23-AUG-2002; 2002WO-GB003892.  
XX  
PR 23-AUG-2001; 2001GB-00020558.  
PR 05-OCT-2001; 2001GB-00024037.  
XX  
PA (OXFO-) OXFORD BIOMEDICA UK LTD.  
XX  
PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;  
PI N-PSDB; ADD19029.  
XX  
DR WPI; 2003-290046/28.  
XX  
PT New substantially purified polypeptide, useful for diagnosing or treating  
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion  
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or  
PT wound healing.  
XX  
PS Claim 1; SEQ ID NO 517; 424pp; English.  
XX  
CC This invention relates to novel human genes and gene product which are  
CC implicated in certain disease states. Compounds which modulate the  
CC proteins of the invention may have cytostatic, anti-inflammatory, the  
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The  
CC sequences of the invention may be useful for gene therapy. The invention  
CC may be useful for diagnosing or treating a hypoxia-regulated condition,  
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,  
CC erythropoiesis, or the biological response to hypoxia conditions  
CC including processes such as glycolysis, gluconeogenesis, glucose  
CC transportation, catecholamine synthesis, iron transport or nitric oxide  
CC synthesis. The disease includes cancer, pre-eclampsia, atherosclerosis  
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis  
CC inflammatory conditions or wound healing. The present sequence is that of  
CC a disease related protein of the invention.  
XX  
SQ Sequence 162 AA;  
XX  
Query Match 77.8%; Score 35; DB 7; Length 162;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 LSYDLFV 8  
DB 18 LSYDLFV 24  
XX  
RESULT 19  
ADJ45973  
ID ADJ45973 standard; protein; 162 AA.  
XX  
AC ADJ45973;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Novel human secreted protein-related protein sequence SegID126.  
XX  
KW secreted protein; upstream regulator; gene therapy; protein purification;  
KW protein synthesis; chromosomal mapping; individual identification;  
KW forensic; hereditary disease; drug reaction; immunoassay;  
KW epitope mapping; vaccine; immune system regulation;  
KW haematopoietic system; tissue growth; reproductive hormone;  
KW cell migration; blood clotting; receptor/ligand interaction;



KW adhesion molecule; assisted drug delivery;  
 KW human glial maturation factor gamma-2; neurite outgrowth;  
 KW neurite resprouting; human.  
 OS Homo sapiens.  
 XX US2003144490-A1.  
 XX  
 XX 31-JUL-2003.  
 XX  
 XX 10-DEC-2002; 2002US-00319763.  
 XX  
 XX 13-NOV-1997; 97US-0066677P.  
 XX 17-DEC-1997; 97US-0069957P.  
 XX 09-FEB-1998; 98US-0074121P.  
 XX 13-APR-1998; 98US-0081563P.  
 XX 10-AUG-1998; 98US-0096116P.  
 XX 04-SEP-1998; 98US-0099273P.  
 XX 13-NOV-1998; 98US-00191997.  
 XX 15-SEP-2000; 2000US-00663600.  
 XX  
 XX (EDMA/) EDWARDS J D M.  
 XX (DUCI/) DUCLETT A.  
 XX (BOUG/) BOUGUELERET L.  
 XX  
 XX Edwards JDM, Ducleert A, Bougueleret L;  
 XX  
 XX WPI; 2003-851788/79.  
 XX N-PSDB; ADJ45926.  
 XX  
 XX New nucleic acid encoding secreted human polypeptides, useful e.g. in  
 XX gene therapy or diagnosis, also encoded proteins, potential therapeutic  
 XX agents.  
 XX  
 XX Example 28; SEQ ID NO 126; 269pp; English.  
 XX  
 CC This invention relates to novel purified isolated polynucleotides which  
 CC comprise a sequence that encodes at least 10 amino acids (aa) from any of  
 CC 48 secreted polypeptide sequences, given in the specification, or  
 CC fragments of polypeptides encoded by human cDNA contained in the  
 CC corresponding deposited clone. The DNA sequences of the invention encode  
 CC secreted proteins (or their fragments) and can be used to  
 CC identify/isolate upstream regulators, potentially useful in gene therapy  
 CC or protein purification, by controlling protein synthesis, as probes for  
 CC chromosomal mapping, identification of individuals, and for diagnosis or  
 CC forensics, for example identifying genes associated with hereditary  
 CC diseases or drug reactions, for recombinant expression of the encoded  
 CC proteins or, where the DNA sequence encodes a signal peptide, for  
 CC directing secretion of heterologous polypeptides. Polypeptides encoded by  
 CC the DNA sequences of the invention can be used to raise antibodies,  
 CC useful for detecting the polypeptide, as (ant)agonists, or for preparing  
 CC anti-idiotypic antibodies, as tags in for example immunoassays, epitope  
 CC mapping or vaccines, also as molecular weight markers, to screen for  
 CC agents with biological activity and as therapeutic agents with,  
 CC potentially, a very wide range of activities, for example regulation of  
 CC the immune or haematopoietic systems, tissue growth, reproductive  
 CC hormones, cell migration, blood clotting or receptor/ligand interaction,  
 CC also as adhesion molecules for assisted drug delivery. A typical isolated  
 CC sequence is human glial maturation factor gamma-2, which stimulates  
 CC neurite outgrowth and resprouting. The present sequence is that of a  
 CC signal and partial mature sequence of a human secreted protein of the  
 CC invention.  
 CC  
 XX  
 XX Sequence 162 AA;  
 XX  
 Query Match 77.8%; Score 35; DB 7; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2 LSYDLFV 8  
 18 LSYDLFV 24

RESULT 20  
 ADJ46067 standard; protein, 162 AA.  
 ID ADJ46067  
 AC ADJ46067;  
 XX  
 XX 06-MAY-2004 (first entry)  
 XX  
 XX Novel human secreted protein-related protein sequence SeqID220.  
 XX  
 KW secreted protein; upstream regulator; gene therapy; protein purification;  
 KW protein synthesis; chromosomal mapping; individual identification;  
 KW forensic; hereditary disease; drug reaction; immunoassay;  
 KW epitope mapping; vaccine; immune system regulation;  
 KW haematopoietic system; tissue growth; reproductive hormone;  
 KW cell migration; blood clotting; receptor/ligand interaction;  
 KW adhesion molecule; assisted drug delivery;  
 KW human glial maturation factor gamma-2; neurite outgrowth;  
 KW neurite resprouting; human.  
 KW  
 XX Homo sapiens.  
 XX  
 XX US2003144490-A1.  
 XX  
 XX 31-JUL-2003.  
 XX  
 XX 10-DEC-2002; 2002US-00319763.  
 XX  
 XX 13-NOV-1997; 97US-0066677P.  
 XX 17-DEC-1997; 97US-0069957P.  
 XX 09-FEB-1998; 98US-0074121P.  
 XX 13-APR-1998; 98US-0081563P.  
 XX 10-AUG-1998; 98US-0096116P.  
 XX 04-SEP-1998; 98US-0099273P.  
 XX 13-NOV-1998; 98US-00191997.  
 XX 15-SEP-2000; 2000US-00663600.  
 XX  
 XX (EDMA/) EDWARDS J D M.  
 XX (DUCI/) DUCLETT A.  
 XX (BOUG/) BOUGUELERET L.  
 XX  
 XX Edwards JDM, Ducleert A, Bougueleret L;  
 XX  
 XX WPI; 2003-851788/79.  
 XX N-PSDB; ADJ46020.  
 XX  
 CC New nucleic acid encoding secreted human polypeptides, useful e.g. in  
 CC gene therapy or diagnosis, also encoded proteins, potential therapeutic  
 CC agents.  
 CC  
 CC Claim 11; SEQ ID NO 220; 269pp; English.  
 CC  
 CC This invention relates to novel purified isolated polynucleotides which  
 CC comprise a sequence that encodes at least 10 amino acids (aa) from any of  
 CC 48 secreted polypeptide sequences, given in the specification, or  
 CC fragments of polypeptides encoded by human cDNA contained in the  
 CC corresponding deposited clone. The DNA sequences of the invention encode  
 CC secreted proteins (or their fragments) and can be used to  
 CC identify/isolate upstream regulators, potentially useful in gene therapy  
 CC or protein purification, by controlling protein synthesis, as probes for  
 CC chromosomal mapping, identification of individuals, and for diagnosis or  
 CC forensics, for example identifying genes associated with hereditary  
 CC diseases or drug reactions, for recombinant expression of the encoded  
 CC proteins or, where the DNA sequence encodes a signal peptide, for  
 CC directing secretion of heterologous polypeptides. Polypeptides encoded by  
 CC the DNA sequences of the invention can be used to raise antibodies,  
 CC useful for detecting the polypeptide, as (ant)agonists, or for preparing  
 CC anti-idiotypic antibodies, as tags in for example immunoassays, epitope  
 CC mapping or vaccines, also as molecular weight markers, to screen for  
 CC agents with biological activity and as therapeutic agents with,  
 CC potentially, a very wide range of activities, for example regulation of  
 CC the immune or haematopoietic systems, tissue growth, reproductive

CC hormones, cell migration, blood clotting or receptor/ligand interaction,  
 CC also as adhesion molecules for assisted drug delivery. A typical isolated  
 CC sequence is human glial maturation factor gamma-2, which stimulates  
 CC neurite outgrowth and resprouting. The present sequence is that of a  
 CC human secreted protein of the invention.

XX Sequence 162 AA;

Query Match 77.8%; Score 35; DB 7; Length 162;

Best Local Similarity 100.0%; Pred.No. 39; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
 |||||  
 18 LSYDLFV 24

Db 18 LSYDLFV 24

RESULT 21

ID ADP19476 standard; protein; 162 AA.

AC ADP19476;

DT 26-AUG-2004 (first entry)

XX Human secreted polypeptide #327.

XX Human; secreted protein; genetic disease.

XX Homo sapiens.

XX US2004110939-A1.

XX 10-JUN-2004.

XX 15-OCT-2001; 2001US-00978360.

XX 17-DEC-1998; 98WO-IB002122.

XX 09-FEB-1999; 99WO-IB000282.

XX 21-JUN-2000; 2000WO-IB000951.

XX 15-SEP-2000; 2000US-0063600.

XX (GSEST ) GENSET SA.

XX Dumas Mline Edwards J, Bougueleret L, Jobert S, Clusel C;

XX Duglert A;

XX WPI; 2004-440404/41.

XX N-PSDB; ADP19071.

XX Claim 2; SEQ ID NO 732; 113bp; English.

XX The invention relates to human cDNA sequences that encode human secreted

XX proteins. The invention also relates to an antibody that specifically

XX binds to a polypeptide of the invention and a method of binding the

XX polypeptide to an antibody. The polynucleotides are useful for expressing

XX the entire secreted proteins which they encode and for distinguishing

XX human tissues and cells from non-human tissues and cells, and for

XX distinguishing between human tissues and cells that do or do not express

XX the polynucleotides comprising the cDNAs. The polynucleotides and

XX polypeptides are useful in forensic procedures or diagnostic procedures

XX to identify individuals with genetic diseases resulting from abnormal

XX expression of the genes corresponding to the cDNAs. The sequences are

XX also useful in gene therapy to control or treat genetic diseases. This

XX sequence represents a human secreted polypeptide of the invention. Note:

XX The sequence data for this patent did not form part of the printed

XX specification but was obtained in electronic format from USPTO at

XX seqdata.uspto.gov/sequence.html.

SQ Sequence 162 AA;

Query Match 77.8%; Score 35; DB 8; Length 162;

Best Local Similarity 100.0%; Pred.No. 39; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
 |||||  
 18 LSYDLFV 24

Db 18 LSYDLFV 24

RESULT 22

ID AAW74871 standard; protein; 163 AA.

AC AAW74871;

DT 25-JUN-1999 (first entry)

XX Human secreted protein encoded by gene 143 clone HBMDW46.

XX Human; secreted protein; testis; tumour; foetal brain tissue;

XX fusion protein; cancer; central nervous system; seizure; diagnosis;

XX neurodegenerative disease.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 163 /label= unknown

XX 11-SEP-1998.

XX 06-MAR-1998; 98WO-US004493.

XX 07-MAR-1997; 97US-0038621P.

XX 07-MAR-1997; 97US-0040161P.

XX 07-MAR-1997; 97US-0040162P.

XX 07-MAR-1997; 97US-0040163P.

XX 07-MAR-1997; 97US-0040333P.

XX 07-MAR-1997; 97US-0040334P.

XX 07-MAR-1997; 97US-0040336P.

XX 07-MAR-1997; 97US-0040626P.

XX 11-APR-1997; 97US-0043311P.

XX 11-APR-1997; 97US-0043312P.

XX 11-APR-1997; 97US-0043313P.

XX 11-APR-1997; 97US-0043314P.

XX 11-APR-1997; 97US-0043315P.

XX 11-APR-1997; 97US-0043568P.

XX 11-APR-1997; 97US-0043569P.

XX 11-APR-1997; 97US-0043576P.

XX 11-APR-1997; 97US-0043578P.

XX 11-APR-1997; 97US-0043580P.

XX 11-APR-1997; 97US-0043669P.

XX 11-APR-1997; 97US-0043670P.

XX 11-APR-1997; 97US-0043671P.

XX 11-APR-1997; 97US-0043672P.

XX 11-APR-1997; 97US-0043674P.

XX 11-APR-1997; 97US-0047492P.

XX 23-MAY-1997; 97US-0047500P.

XX 23-MAY-1997; 97US-0047501P.

XX 23-MAY-1997; 97US-0047502P.

XX 23-MAY-1997; 97US-0047503P.

XX 23-MAY-1997; 97US-0047581P.

XX 23-MAY-1997; 97US-0047582P.

XX 23-MAY-1997; 97US-0047583P.

XX 23-MAY-1997; 97US-0047584P.

XX 23-MAY-1997; 97US-0047585P.

XX 23-MAY-1997; 97US-0047586P.

XX 23-MAY-1997; 97US-0047587P.

XX 23-MAY-1997; 97US-0047588P.

PR	23-MAY-1997	97US-0047589P
PR	23-MAY-1997	97US-0047590P
PR	23-MAY-1997	97US-0047592P
PR	23-MAY-1997	97US-0047593P
PR	23-MAY-1997	97US-0047594P
PR	23-MAY-1997	97US-0047595P
PR	23-MAY-1997	97US-0047596P
PR	23-MAY-1997	97US-0047597P
PR	23-MAY-1997	97US-0047598P
PR	23-MAY-1997	97US-0047599P
PR	23-MAY-1997	97US-0047600P
PR	23-MAY-1997	97US-0047601P
PR	23-MAY-1997	97US-0047612P
PR	23-MAY-1997	97US-0047613P
PR	23-MAY-1997	97US-0047614P
PR	23-MAY-1997	97US-0047615P
PR	23-MAY-1997	97US-0047617P
PR	23-MAY-1997	97US-0047618P
PR	23-MAY-1997	97US-0047632P
PR	23-MAY-1997	97US-0047633P
PR	06-JUN-1997	97US-0048964P
PR	06-JUN-1997	97US-0048974P
PR	13-JUN-1997	97US-0049610P
PR	08-JUL-1997	97US-0051926P
PR	16-JUL-1997	97US-0052874P
PR	18-AUG-1997	97US-0055724P
PR	22-AUG-1997	97US-0056630P
PR	22-AUG-1997	97US-0056631P
PR	22-AUG-1997	97US-0056632P
PR	22-AUG-1997	97US-0056636P
PR	22-AUG-1997	97US-0056637P
PR	22-AUG-1997	97US-0056662P
PR	22-AUG-1997	97US-0056674P
PR	22-AUG-1997	97US-0056675P
PR	22-AUG-1997	97US-0056684P
PR	22-AUG-1997	97US-0056682P
PR	22-AUG-1997	97US-0056677P
PR	22-AUG-1997	97US-0056678P
PR	22-AUG-1997	97US-0056680P
PR	22-AUG-1997	97US-0056812P
PR	22-AUG-1997	97US-0056822P
PR	22-AUG-1997	97US-0056844P
PR	22-AUG-1997	97US-0056866P
PR	22-AUG-1997	97US-0056876P
PR	22-AUG-1997	97US-0056880P
PR	22-AUG-1997	97US-0056881P
PR	22-AUG-1997	97US-0056882P
PR	22-AUG-1997	97US-0056892P
PR	22-AUG-1997	97US-0056893P
PR	22-AUG-1997	97US-0056894P
PR	22-AUG-1997	97US-0056903P
PR	22-AUG-1997	97US-0056908P
PR	22-AUG-1997	97US-0056909P
PR	22-AUG-1997	97US-0056910P
PR	22-AUG-1997	97US-0056911P
PR	05-SEP-1997	97US-0057650P
PR	05-SEP-1997	97US-0057659P
PR	05-SEP-1997	97US-0057761P
PR	12-SEP-1997	97US-0058785P
PR	02-OCT-1997	97US-0061060P
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PI	Ruben SM, Rosen CA, Fischer CU, Soppet DR, Carter KC,	
PI	Beharati DP, Ehdress GA, Yu G, Ni J, Feng P, Young PE, Greene JW,	
PI	Ferrie AM, Duan R, Hu Y, Florence KA, Olsen HS, Ebner R, Brewer LA,	
PI	Moore PA, Shi Y, Latleur DW, Li Y, Zeng Z, Kraw H,	
PR	WPI; 1998-506364/43.	

DR	N-PSDB; AAV59653.
XX	
PT	New isolated human genes and the secreted polypeptide(s) they encode -
PT	useful for diagnosis and treatment of e.g. cancers, neurological
PT	disorders, immune diseases, inflammation or blood disorders.
XX	
PS	Claim 1, Page 626-627; 721np; English.
XX	
CC	This sequence represents a secreted human protein encoded by the nucleic
CC	acid molecule designated Gene IBMDM46 from the human cDNA clone HTTEZ21
CC	(deposited as clone ATCC 97902 and ATCC 209048). The gene can be used to
CC	generate fusion proteins by linking to the gene to a human immunoglobulin
CC	Fc portion (e.g. AAV59502) for increasing the stability of the fused
CC	protein as compared to the human protein only. The invention relates to
CC	18 novel genes and their fragments (nucleic acid sequences: AAV59511-
CC	V59812; amino acid sequences AAW74731-W75026) which are useful for
CC	preventing, treating or ameliorating medical conditions e.g. by protein
CC	or gene therapy. Also, pathological conditions in a sample or by
CC	determining the amount of the new polypeptides can be diagnosed by
CC	determining the presence of mutations in the new polynucleotides.
CC	Specific uses are described for each of the 16 polynucleotides, based on
CC	which tissues they are most highly expressed in (see AAV59511 for
CC	described uses)
XX	
SQ	Sequence 163 AA;
Query Match	77.8%; Score 35; DB 2; Length 163;
Best Local Similarity	100.0%; Pred. No. 40;
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	2 LSYDLFV 8         
Dd	18 LSYDLFV 24
RESULT 23	
ABG95322	
ID	ABG95322 standard; protein; 163 AA.
XX	
AC	ABG95322;
XX	
DT	15-JAN-2003 (first entry)
XX	
DE	Human novel secreted protein #143.
KW	Human; secreted protein; autoimmune disease; chemotaxis;
KW	rheumatoid arthritis; hyperproliferative disorder; breast neoplasm;
KW	liver neoplasm cardiovascular disorder; cardiac arrest; skin aging;
KW	cerebrovascular disorder; cerebral ischaemia; angiogenesis; sunburn;
KW	nervous system disorders; Alzheimer's disease; infection;
KW	ocular disorder; corneal infection; wound healing; tissue regeneration;
KW	epithelial cell proliferation; organ transplantation; food additive;
KW	preservative; nutritional.
OS	Homo sapiens.
XX	
XX	
PN	US6420526-B1.
XX	
PD	16-JUL-2002.
XX	
PF	08-SEP-1998; 98US-00149476.
XX	
PR	07-MAR-1997; 97US-0038621P.
PR	07-MAR-1997; 97US-0040161P.
PR	07-MAR-1997; 97US-0040162P.
PR	07-MAR-1997; 97US-0040163P.
PR	07-MAR-1997; 97US-0040334P.
PR	07-MAR-1997; 97US-0040334P.
PR	07-MAR-1997; 97US-0040336P.
PR	07-MAR-1997; 97US-0040626P.
PR	11-APR-1997; 97US-0043311P.
PR	11-APR-1997; 97US-0043312P.
PR	11-APR-1997; 97US-0043313P.

PR 11-APR-1997; 97US-0043314P.  
 PR 11-APR-1997; 97US-0043315P.  
 PR 11-APR-1997; 97US-0043568P.  
 PR 11-APR-1997; 97US-0043569P.  
 PR 11-APR-1997; 97US-0043576P.  
 PR 11-APR-1997; 97US-0043578P.  
 PR 11-APR-1997; 97US-0043580P.  
 PR 11-APR-1997; 97US-0043669P.  
 PR 11-APR-1997; 97US-0043670P.  
 PR 11-APR-1997; 97US-0043671P.  
 PR 11-APR-1997; 97US-0043672P.  
 PR 11-APR-1997; 97US-0043674P.  
 PR 23-MAY-1997; 97US-0047492P.  
 PR 23-MAY-1997; 97US-0047500P.  
 PR 23-MAY-1997; 97US-0047501P.  
 PR 23-MAY-1997; 97US-0047502P.  
 PR 23-MAY-1997; 97US-0047503P.  
 PR 23-MAY-1997; 97US-0047581P.  
 PR 23-MAY-1997; 97US-0047582P.  
 PR 23-MAY-1997; 97US-0047583P.  
 PR 23-MAY-1997; 97US-0047584P.  
 PR 23-MAY-1997; 97US-0047585P.  
 PR 23-MAY-1997; 97US-0047586P.  
 PR 23-MAY-1997; 97US-0047587P.  
 PR 23-MAY-1997; 97US-0047588P.  
 PR 23-MAY-1997; 97US-0047589P.  
 PR 23-MAY-1997; 97US-0047590P.  
 PR 23-MAY-1997; 97US-0047592P.  
 PR 23-MAY-1997; 97US-0047593P.  
 PR 23-MAY-1997; 97US-0047594P.  
 PR 23-MAY-1997; 97US-0047595P.  
 PR 23-MAY-1997; 97US-0047596P.  
 PR 23-MAY-1997; 97US-0047597P.  
 PR 23-MAY-1997; 97US-0047598P.  
 PR 23-MAY-1997; 97US-0047599P.  
 PR 23-MAY-1997; 97US-0047600P.  
 PR 23-MAY-1997; 97US-0047601P.  
 PR 23-MAY-1997; 97US-0047612P.  
 PR 23-MAY-1997; 97US-0047613P.  
 PR 23-MAY-1997; 97US-0047614P.  
 PR 23-MAY-1997; 97US-0047615P.  
 PR 23-MAY-1997; 97US-0047617P.  
 PR 23-MAY-1997; 97US-0047618P.  
 PR 23-MAY-1997; 97US-0047632P.  
 PR 23-MAY-1997; 97US-0047633P.  
 PR 06-JUN-1997; 97US-0048964P.  
 PR 06-JUN-1997; 97US-0048974P.  
 PR 13-JUN-1997; 97US-0049610P.  
 PR 08-JUL-1997; 97US-0051926P.  
 PR 16-JUL-1997; 97US-0052874P.  
 PR 18-AUG-1997; 97US-0055724P.  
 PR 22-AUG-1997; 97US-0056630P.  
 PR 22-AUG-1997; 97US-0056631P.  
 PR 22-AUG-1997; 97US-0056632P.  
 PR 22-AUG-1997; 97US-0056636P.  
 PR 22-AUG-1997; 97US-0056637P.  
 PR 22-AUG-1997; 97US-0056662P.  
 PR 22-AUG-1997; 97US-0056664P.  
 PR 22-AUG-1997; 97US-0056684P.  
 PR 22-AUG-1997; 97US-0056862P.  
 PR 22-AUG-1997; 97US-0056864P.  
 PR 22-AUG-1997; 97US-0056872P.  
 PR 22-AUG-1997; 97US-0056874P.  
 PR 22-AUG-1997; 97US-0056875P.  
 PR 22-AUG-1997; 97US-0056876P.  
 PR 22-AUG-1997; 97US-0056877P.  
 PR 22-AUG-1997; 97US-0056879P.  
 PR 22-AUG-1997; 97US-0056880P.  
 PR 22-AUG-1997; 97US-0056881P.  
 PR 22-AUG-1997; 97US-0056882P.  
 PR 22-AUG-1997; 97US-0056884P.  
 PR 22-AUG-1997; 97US-0056886P.

PR 22-AUG-1997; 97US-0056887P.  
 PR 22-AUG-1997; 97US-0056888P.  
 PR 22-AUG-1997; 97US-0056889P.  
 PR 22-AUG-1997; 97US-0056892P.  
 PR 22-AUG-1997; 97US-0056893P.  
 PR 22-AUG-1997; 97US-0056894P.  
 PR 22-AUG-1997; 97US-0056903P.  
 PR 22-AUG-1997; 97US-0056908P.  
 PR 22-AUG-1997; 97US-0056909P.  
 PR 22-AUG-1997; 97US-0056910P.  
 PR 22-AUG-1997; 97US-0056911P.  
 PR 05-SEP-1997; 97US-0057650P.  
 PR 05-SEP-1997; 97US-0057659P.  
 PR 05-SEP-1997; 97US-0057761P.  
 PR 12-SEP-1997; 97US-0058785P.  
 PR 02-OCT-1997; 97US-0061060P.  
 PR 06-MAR-1998; 98WO-US004493.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC, Greene JM;  
 PI Bednarik DR, Endress GA, Yu G, Ni J, Feng P, Young PE, Brewer LA;  
 PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA;  
 PI Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
 XX  
 XX WPI; 2002-634796/68.  
 DR N-PSDB; ABS73640.  
 XX  
 PT New isolated human secreted protein for diagnosing, preventing, treating  
 PT or ameliorating medical conditions and used as a food additive or  
 PT preservative.  
 PT  
 XX  
 XX Example 1; SEQ ID NO 462; 129pp; English.  
 PS  
 CC The invention relates to an isolated protein that is one of 186 human  
 CC secreted proteins, given in the specification. The protein is used in a  
 CC sequences also given in the specification. The protein is used in a  
 CC pharmaceutical composition used to prevent, treat or ameliorate a medical  
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
 CC chickens or sheep. Disorders which are diagnosed or treated include  
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angioneuromatosis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to  
 CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,  
 CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
 CC components. The present sequence represents one of the novel human  
 CC secreted proteins of the invention. Note: This sequence did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocId=6420526B1  
 XX  
 SQ Sequence 163 AA:  
 Query Match 77.8%; Score 35; DB 5; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LSYDLFV 8  
 Db 18 LSYDLFV 24  
 RESULT 24  
 ABO34516  
 ID ABO34516 standard; protein; 163 AA.  
 XX  
 AC ABO34516;



PA (FLOR/) FLORENCE K. A.  
PA (OLSE/) OLSEN H. S.  
PA (RISC/) RISCHER C. L.  
PA (EBNE/) EBNER R.  
PA (BREW/) BREWER L. A.  
PA (MOOR/) MOORE P. A.  
PA (SHIY/) SHI Y.  
PA (LAFLE/) LAFLEUR D. W.  
PA (LIYV/) LI Y.  
PA (ZENG/) ZENG Z.  
PA (KYAW/) KYAW H.  
XX  
XX Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednarik DP,  
PI Endress GA, Yu G, Ni J, Feng P, Young PB, Greene JM, Ferrie AM;  
PI Duan DR, Hu J, Florence KA, Olsen HS, Fischer CL, Ebner R;  
PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
XX WPI; 2003-521800/49.  
DR N-PSDB; ACD82783.  
XX  
XX New genes and its encoded prostate cancer antigen proteins, useful for  
PT preventing, treating, ameliorating or diagnosing e.g. prostate cancers,  
PT lymphic hypoplasia, multiple sclerosis, AIDS, angina pectoris or cerebral  
PT ischemia.  
XX  
XX Claim 3; SEQ ID NO 462; 260pp; English.  
XX  
XX The present invention relates to the isolation of novel human secreted  
CC proteins and the polynucleotide sequences encoding them. The invention  
CC also discloses vectors, host cells, antibodies, and recombinant methods  
CC for producing human secreted proteins. The polypeptide and polynucleotide  
CC sequences for the secreted proteins are useful for preventing, treating,  
CC ameliorating or diagnosing medical conditions such as hyperproliferative  
CC disorders (e.g. leukemia or breast cancers), wounds, reproductive  
CC disorders, blood-related disorders (e.g. haemophilia or  
CC thrombocytopaenia), immunodeficiencies (e.g. Wiskott-Aldrich syndrome or  
CC lymphic hypoplasia), autoimmune disorders (e.g. graft-versus-host disease,  
CC multiple sclerosis or Hashimoto's thyroiditis), allergies (e.g. asthma),  
CC viral or bacterial or fungal infections (e.g. AIDS or sepsis), renal  
CC disorders (e.g. kidney failure), cardiovascular disorders (e.g. angina  
CC pectoris, cerebral ischaemia or congenital heart defects), respiratory  
CC disorders, neurological disorders (e.g. Alzheimer's disease or  
CC Parkinson's disease), and inflammations (e.g. Crohn's disease). The  
CC polynucleotide or polypeptide may also be used as vaccine adjuvants.  
CC AB034374-AB034815 represent human secreted proteins or their fragments.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO web site at [seqdata.uspto.gov/psipdsidentry.html](http://seqdata.uspto.gov/psipdsidentry.html)  
CC  
XX  
SQ Sequence 163 AA;  
  
Query Match 77.8%; Score 35; DB 6; Length 163;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 LSYDLFV 8  
Db 18 LSYDLFV 24  
  
RESULT 25  
AD123177  
ID AD123177 standard; protein; 163 AA.  
XX  
XX AD123177;  
AC  
XX 22-APR-2004 (first entry)  
DT  
XX  
DE Novel human secreted protein seq id 462.  
XX  
KW cytostatic; gene therapy; cancer; human; secreted protein.  
XX  
OS Homo sapiens.

XX  
XX US2003175858-A1.  
XX  
XX 18-SEP-2003.  
PD  
XX  
XX 18-JUN-2001; 2001US-00882171.  
PF  
XX  
XX 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040336P.  
PR 07-MAR-1997; 97US-0040626P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043368P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
PR 11-APR-1997; 97US-004369P.  
PR 11-APR-1997; 97US-0043670P.  
PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 11-APR-1997; 97US-0043674P.  
PR 23-MAY-1997; 97US-0047492P.  
PR 23-MAY-1997; 97US-0047501P.  
PR 23-MAY-1997; 97US-0047501P.  
PR 23-MAY-1997; 97US-0047502P.  
PR 23-MAY-1997; 97US-0047503P.  
PR 23-MAY-1997; 97US-0047581P.  
PR 23-MAY-1997; 97US-0047582P.  
PR 23-MAY-1997; 97US-0047583P.  
PR 23-MAY-1997; 97US-0047584P.  
PR 23-MAY-1997; 97US-0047585P.  
PR 23-MAY-1997; 97US-0047586P.  
PR 23-MAY-1997; 97US-0047587P.  
PR 23-MAY-1997; 97US-0047588P.  
PR 23-MAY-1997; 97US-0047589P.  
PR 23-MAY-1997; 97US-0047590P.  
PR 23-MAY-1997; 97US-0047592P.  
PR 23-MAY-1997; 97US-0047593P.  
PR 23-MAY-1997; 97US-0047594P.  
PR 23-MAY-1997; 97US-0047595P.  
PR 23-MAY-1997; 97US-0047596P.  
PR 23-MAY-1997; 97US-0047597P.  
PR 23-MAY-1997; 97US-0047598P.  
PR 23-MAY-1997; 97US-0047599P.  
PR 23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047617P.  
PR 23-MAY-1997; 97US-0047618P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047632P.  
PR 23-MAY-1997; 97US-0047633P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 06-JUN-1997; 97US-0048974P.  
PR 13-JUN-1997; 97US-0049610P.  
PR 08-JUL-1997; 97US-0051926P.  
PR 16-JUL-1997; 97US-0052874P.  
PR 18-AUG-1997; 97US-0055724P.  
PR 22-AUG-1997; 97US-0056630P.  
PR 22-AUG-1997; 97US-0056631P.  
PR 22-AUG-1997; 97US-0056632P.  
PR 22-AUG-1997; 97US-0056636P.  
PR 22-AUG-1997; 97US-0056637P.

PR 22-AUG-1997; 97US-0056662P.  
PR 22-AUG-1997; 97US-0056664P.  
PR 22-AUG-1997; 97US-0056845P.  
PR 22-AUG-1997; 97US-0056862P.  
PR 22-AUG-1997; 97US-0056866P.  
PR 22-AUG-1997; 97US-0056872P.  
PR 22-AUG-1997; 97US-0056875P.  
PR 22-AUG-1997; 97US-0056876P.  
PR 22-AUG-1997; 97US-0056877P.  
PR 22-AUG-1997; 97US-0056878P.  
PR 22-AUG-1997; 97US-0056879P.  
PR 22-AUG-1997; 97US-0056880P.  
PR 22-AUG-1997; 97US-0056881P.  
PR 22-AUG-1997; 97US-0056882P.  
PR 22-AUG-1997; 97US-0056884P.  
PR 22-AUG-1997; 97US-0056886P.  
PR 22-AUG-1997; 97US-0056887P.  
PR 22-AUG-1997; 97US-0056888P.  
PR 22-AUG-1997; 97US-0056889P.  
PR 22-AUG-1997; 97US-0056892P.  
PR 22-AUG-1997; 97US-0056893P.  
PR 22-AUG-1997; 97US-0056894P.  
PR 22-AUG-1997; 97US-0056903P.  
PR 22-AUG-1997; 97US-0056908P.  
PR 22-AUG-1997; 97US-0056909P.  
PR 22-AUG-1997; 97US-0056910P.  
PR 22-AUG-1997; 97US-0056911P.  
PR 05-SEP-1997; 97US-0057650P.  
PR 05-SEP-1997; 97US-0057663P.  
PR 05-SEP-1997; 97US-0057761P.  
PR 12-SEP-1997; 97US-0058785P.  
PR 09-OCT-1997; 97US-0061660P.  
PR 06-MAR-1998; 98MO-US004493.  
PR 08-SEP-1998; 98US-00149476.  
PR 17-MAR-2000; 2000US-0180068P.  
PR 16-MAR-2001; 2001US-00809391P.

XX (RUBE/) RUBEN S. M.  
PA (ROSE/) ROSEN C. A.  
PA (SOP/) SOPPET D. R.  
PA (CART/) CARTER K. C.  
PA (BEDN/) BEDNARIK D. P.  
PA (ENDR/) ENDRESS G. A.  
PA (YUGG/) YU G.  
PA (NIJ/) NI J.  
PA (FENG/) FENG P. P.  
PA (YOUN/) YOUNG P. E.  
PA (GREE/) GREENE J. M.  
PA (FERR/) FERRIE A. M.  
PA (DUAN/) DUAN D. R.  
PA (HUJ/) HU J.  
PA (FLO/) FLORENCE K. A.  
PA (OLSE/) OLSEN H. S.  
PA (FISC/) FISCHER C. L.  
PA (EBNE/) EBNER R.  
PA (BREW/) BREWER L. A.  
PA (MOOR/) MOORE P. A.  
PA (SHIY/) SHI Y.  
PA (LAFLE/) LAFLEUR D. W.  
PA (LIY/) LI Y.  
PA (ZENG/) ZENG Z.  
PA (KYAW/) KYAW H.  
XX  
PI Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednarik DP,  
PI Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM, Ferrie AM,  
PI Duan DR, Hu J, Fischer CA, Olsen HS, Ebner R,  
PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
XX WPI: 2003-898535/82.  
DR N-PSDB; ADI22868.  
XX  
PT New nucleic acid molecule, useful for preparing a medicament for

PT diagnosing, preventing, treating or ameliorating a medical condition  
e.g., cancer.  
XX  
XX  
PS Claim 11; SEQ ID NO 462; 256pp; English.

CC The invention describes an isolated nucleic acid comprising a sequence  
CC having 95 % identity with: a polynucleotide fragment of a sequence not  
CC given in the specification, or its allelic variant; a polynucleotide  
CC fragment of the cDNA sequence; a polynucleotide sequence encoding a  
CC polypeptide, or its fragment, domain, epitope or species homologue; or a  
CC polynucleotide that hybridises under stringent conditions to any one of  
CC the sequences of (a)-(c). The nucleic acid is useful for preparing a  
CC medicament for diagnosing, preventing, treating or ameliorating a medical  
CC condition e.g., cancer. The is the amino acid sequence of a novel human  
CC secreted protein of the invention.

SO Sequence 163 AA;

Query Match 77.8%; Score 35; DB 7; Length 163;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
Db 18 LSYDLFV 24

RESULT 26  
ID ADH74179 standard; protein; 163 AA.

XX ADH74179;

AC 25-MAR-2004 (first entry)

DT Human secreted protein #143.

DE human; secreted protein; cancer; haematopoietic disorder;

KW endocrine disorder; immune system disease; inflammatory disorder.

XX Homo sapiens.

XX US2003225248-A1.

PD 04-DEC-2003.

XX 10-JUN-2002; 2002US-00164861.

XX 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040161P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040336P.  
PR 07-MAR-1997; 97US-0040626P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043568P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
PR 11-APR-1997; 97US-0043669P.  
PR 11-APR-1997; 97US-0043670P.  
PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 11-APR-1997; 97US-0043674P.  
PR 23-MAY-1997; 97US-0047492P.  
PR 23-MAY-1997; 97US-0047500P.

PR 23-MAY-1997; 97US-0047501P.  
 PR 23-MAY-1997; 97US-0047502P.  
 PR 23-MAY-1997; 97US-0047503P.  
 PR 23-MAY-1997; 97US-0047581P.  
 PR 23-MAY-1997; 97US-0047582P.  
 PR 23-MAY-1997; 97US-0047583P.  
 PR 23-MAY-1997; 97US-0047584P.  
 PR 23-MAY-1997; 97US-0047585P.  
 PR 23-MAY-1997; 97US-0047586P.  
 PR 23-MAY-1997; 97US-0047587P.  
 PR 23-MAY-1997; 97US-0047588P.  
 PR 23-MAY-1997; 97US-0047589P.  
 PR 23-MAY-1997; 97US-0047590P.  
 PR 23-MAY-1997; 97US-0047592P.  
 PR 23-MAY-1997; 97US-0047593P.  
 PR 23-MAY-1997; 97US-0047594P.  
 PR 23-MAY-1997; 97US-0047595P.  
 PR 23-MAY-1997; 97US-0047596P.  
 PR 23-MAY-1997; 97US-0047597P.  
 PR 23-MAY-1997; 97US-0047598P.  
 PR 23-MAY-1997; 97US-0047599P.  
 PR 23-MAY-1997; 97US-0047600P.  
 PR 23-MAY-1997; 97US-0047601P.  
 PR 23-MAY-1997; 97US-0047612P.  
 PR 23-MAY-1997; 97US-0047613P.  
 PR 23-MAY-1997; 97US-0047614P.  
 PR 23-MAY-1997; 97US-0047615P.  
 PR 23-MAY-1997; 97US-0047617P.  
 PR 23-MAY-1997; 97US-0047618P.  
 PR 23-MAY-1997; 97US-0047632P.  
 PR 23-MAY-1997; 97US-0047633P.  
 PR 06-JUN-1997; 97US-0048964P.  
 PR 06-JUN-1997; 97US-0048974P.  
 PR 13-JUN-1997; 97US-0049610P.  
 PR 08-JUL-1997; 97US-0051926P.  
 PR 16-JUL-1997; 97US-0052874P.  
 PR 18-AUG-1997; 97US-0055724P.  
 PR 22-AUG-1997; 97US-0056630P.  
 PR 22-AUG-1997; 97US-0056631P.  
 PR 22-AUG-1997; 97US-0056632P.  
 PR 22-AUG-1997; 97US-0056636P.  
 PR 22-AUG-1997; 97US-0056637P.  
 PR 22-AUG-1997; 97US-0056662P.  
 PR 22-AUG-1997; 97US-0056664P.  
 PR 22-AUG-1997; 97US-0056845P.  
 PR 22-AUG-1997; 97US-0056862P.  
 PR 22-AUG-1997; 97US-0056864P.  
 PR 22-AUG-1997; 97US-0056872P.  
 PR 22-AUG-1997; 97US-0056874P.  
 PR 22-AUG-1997; 97US-0056875P.  
 PR 22-AUG-1997; 97US-0056876P.  
 PR 22-AUG-1997; 97US-0056877P.  
 PR 22-AUG-1997; 97US-0056878P.  
 PR 22-AUG-1997; 97US-0056879P.  
 PR 22-AUG-1997; 97US-0056880P.  
 PR 22-AUG-1997; 97US-0056881P.  
 PR 22-AUG-1997; 97US-0056882P.  
 PR 22-AUG-1997; 97US-0056884P.  
 PR 22-AUG-1997; 97US-0056886P.  
 PR 22-AUG-1997; 97US-0056887P.  
 PR 22-AUG-1997; 97US-0056888P.  
 PR 22-AUG-1997; 97US-0056889P.  
 PR 22-AUG-1997; 97US-0056892P.  
 PR 22-AUG-1997; 97US-0056893P.  
 PR 22-AUG-1997; 97US-0056894P.  
 PR 22-AUG-1997; 97US-0056903P.  
 PR 22-AUG-1997; 97US-0056908P.  
 PR 22-AUG-1997; 97US-0056909P.  
 PR 22-AUG-1997; 97US-0056910P.  
 PR 22-AUG-1997; 97US-0056911P.  
 PR 05-SEP-1997; 97US-0057650P.  
 PR 05-SEP-1997; 97US-0057669P.  
 PR 05-SEP-1997; 97US-0057761P.

PR 12-SEP-1997; 97US-0058785P.  
 PR 02-OCT-1997; 97US-0061060P.  
 PR 06-MAR-1998; 98WO-US004493.  
 PR 08-SEP-1998; 98US-00149476.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednarik DP, Ferrie AM;  
 PI Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM, Fischer CL, Ebner R;  
 PI Duan R, Hu J, Florence KA, Olsen HS, Fischer CL, Ebner R;  
 PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
 XX  
 DR MPI, 2004-131264/13.  
 DR N-PSDB; ADH73870.  
 XX  
 PT Isolated nucleic acid molecules encoding human secreted proteins, useful  
 PT for preventing, diagnosing and treating disorders associated with  
 PT aberrant expression and activity.  
 XX  
 PS Claim 11; SEQ ID NO 462, 142pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules and the human  
 CC secreted proteins (SPs) they encode. The proteins and nucleic acids may  
 CC be used in the prevention, diagnosis and treatment of diseases associated  
 CC with inappropriate SP expression e.g. cancer, hematopoietic disorders,  
 CC endocrine disorders, diseases of the immune system, inflammatory  
 CC disorders and many others. Full details of disorders that may be  
 CC prevented, diagnosed and/or treated by the above methods are given in the  
 CC specification. The nucleic acid molecules may be used to produce their  
 CC proteins. The nucleic acid and its complementary sequences may also be  
 CC used as DNA probes in diagnostic assays to detect and quantitate the  
 CC presence of similar nucleic acids in samples, and therefore which  
 CC patients may be in need of restorative therapy. The SPs may also be used  
 CC as antigens in the production of antibodies against the proteins and in  
 CC assays to identify modulators of SP expression and activity. The anti-SP  
 CC antibodies and antagonists may also be used to down regulate expression  
 CC and activity. The anti-SP antibodies may also be used as diagnostic  
 CC agents for detecting the presence of the proteins in samples (e.g. by  
 CC enzyme linked immunosorbent assay (ELISA)). The present sequence  
 CC represents the amino acid sequence of a human secreted protein.  
 CC  
 SQ Sequence 163 AA;  
 QY  
 Db 2 LSYDLFV 8  
 18 LSYDLFV 24  
 RESULT 27  
 ABB97496  
 ID ABB97496 standard; protein; 177 AA.  
 XX  
 AC ABB97496;  
 DT 27-JUN-2002 (first entry)  
 XX  
 DE Novel human protein SEQ ID NO: 764.  
 XX  
 KW Human; antianaemic; vulnery; antiinflammatory; immunomodulator;  
 KW antifertility; cerebroprotective; cytosolic; rheumatic; gene therapy;  
 KW neuroprotective; antiparkinsonian; protein therapy; EST;  
 XX expressed sequence tag.  
 OS Homo sapiens.  
 XX  
 PN WO200222660-A2.  
 XX  
 PD 21-MAR-2002.  
 XX



PF 10-SEP-2001; 2001WO-US026015.  
XX  
PR 11-SEP-2000; 2000US-00659671.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT,  
XX WPI; 2002-292408/33.  
DR N-PSDB; ABN32682.  
XX  
XX An isolated polynucleotide for treating diseases associated with its  
PT encoded polypeptide such as cancer and multiple sclerosis.  
XX  
PS Claim 20; SEQ ID NO 764; 509pp; English.  
XX  
CC The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke  
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.  
CC rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a protein of the invention  
XX  
SQ Sequence 177 AA;  
XX  
Query Match 77.8%; Score 35; DB 5; Length 177;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 LSYDLFV 8  
Db 18 LSYDLFV 24  
XX  
RESULT 28  
AAE25673  
ID AAE25673 standard; protein; 203 AA.  
XX  
XX AAE25673;  
XX  
DT 29-AUG-2003 (revised)  
DT 04-NOV-2002 (first entry)  
XX  
DE Bacillus stearothermophilus BstYI endonuclease protein.  
XX  
KW Recombinant DNA; BstYI restriction endonuclease; BstY methylase;  
KW purification; enzyme.  
XX  
OS Geobacillus stearothermophilus.  
XX  
PN US6403354-B1.  
XX  
PD 11-JUN-2002.  
XX  
PF 19-JAN-2001; 2001US-00766055.  
XX  
PR 19-JAN-2001; 2001US-00766055.  
XX  
PA (NEME ) NEW ENGLAND BIOLABS INC.  
XX  
PI Xu S, Samuelson J, Pelletier J, Sibley M, Wilson GG;  
XX WPI; 2002-536039/57.  
DR N-PSDB; AAD42119.  
XX  
XX Novel DNA encoding BstYI restriction endonuclease and BstY methylase,  
PT useful for producing recombinant BstY restriction endonuclease.  
XX

PS Disclosure; Col 19-20; 18pp; English.  
XX  
XX The invention relates to recombinant DNA encoding BstYI restriction  
CC endonuclease and BstY methylase. The invention also relates to methods  
CC for purification of the recombinant BstYI restriction endonuclease and  
CC BstY methylase. Host cell transformed by the recombinant DNA of the  
CC invention is useful for producing recombinant BstY restriction  
CC endonuclease. The DNA of the invention is an useful tool for creating  
CC recombinant molecules in the laboratory. The present sequence is Bacillus  
CC stearothermophilus BstYI endonuclease protein. (Updated on 29-AUG-2003 to  
CC standardise OS field)  
XX  
XX  
SQ Sequence 203 AA;  
XX  
Query Match 77.8%; Score 35; DB 5; Length 203;  
Best Local Similarity 75.0%; Pred. No. 50;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FLSYDLFV 8  
Db 136 FVAYDLFV 143  
XX  
RESULT 29  
ABR62959  
ID ABR62959 standard; protein; 203 AA.  
XX  
XX ABR62959;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE BstYI restriction endonuclease.  
XX  
KW BstYI; restriction endonuclease; enzyme.  
XX  
OS Bacillus stearothermophilus.  
XX  
PN WO2003060152-A2.  
XX  
PD 24-JUL-2003.  
XX  
PF 09-JAN-2003; 2003WO-US000542.  
XX  
PR 10-JAN-2002; 2002US-0347403P.  
XX  
PA (NEME ) NEW ENGLAND BIOLABS INC.  
XX  
PI Samuelson JC, Xu S;  
XX WPI; 2003-587289/55.  
DR N-PSDB; ACP79289.  
XX  
PT Altering an endonuclease recognition site specificity, comprises  
PT subjecting a mutagenized endonuclease gene library to a genetic selection  
PT in prokaryotic host cells expressing one or more non-cognate DNA  
XX methyltransferases.  
XX  
PS Disclosure; Fig 2; 51pp; English.  
XX  
XX The present sequence is the protein sequence of the BstYI restriction  
CC endonuclease of Bacillus stearothermophilus Y406. BstYI was used in an  
CC example of the method of the invention, which involves altering the DNA  
CC recognition and cleavage characteristics of an endonuclease. The method  
CC can be used without prior knowledge of the endonuclease's three-  
CC dimensional structure and/or the amino acid residues responsible for  
CC activity and/or specificity. It involves subjected a mutagenized  
CC endonuclease gene library to a genetic selection in prokaryotic cells  
CC expressing one or more non-cognate DNA methyltransferases. Genetic  
CC selection selects for viable cells in the population, and viable cells  
CC that express as active mutated endonuclease with an altered recognition  
CC site specificity are identified. A claimed modified BstYI enzyme has the  
CC recognition site 5'-AGATCT-3', altered from 5'-RGATCY-3' for the native  
CC enzyme

XX  
SQ Sequence 203 AA;

Query Match 77.8%; Score 35; DB 7; Length 203;  
Best Local Similarity 75.0%; Pred. No. 50;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLSTYDLFV 8  
|:|:|:|:|  
Db 136 FVAVDLFV 143

RESULT 30  
AA008931

ID AA008931 standard; protein; 51 AA.

AC AA008931;

DT 06-NOV-2001 (first entry)

DE Human polypeptide SEQ ID NO 22823.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorders; arthritis; inflammation.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PE 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

DR N-PSDB; AAI88862.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing

PT and treating e.g. leukemia, inflammation and immune disorders.

PS Claim 20; SEQ ID NO 22823; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and

CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 51 AA;

SQ

Query Match 75.6%; Score 34; DB 4; Length 51;

Best Local Similarity 66.7%; Pred. No. 18;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 FLSTYDLFV 9  
|:|:|:|:|  
Db 42 FLPTDIFTV 50

RESULT 31

ABP64088

ID ABP64088 standard; protein; 55 AA.

AC ABP64088;

DT 04-NOV-2002 (first entry)

DE Human ORF458.

XX Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;

KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;

KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;

KW cancer; cardiovascular disease; allergy; autoimmune disease;

KW wound healing; blood coagulation disorder; inflammatory disorder.

OS Homo sapiens.

PN US2002082206-A1.

PD 27-JUN-2002.

PE 30-MAY-2001; 2001US-00867550.

XX 30-MAY-2000; 2000US-0208427P.

XX (LEAC/) LEACH M D.

PA (MEHR/) MEHRABAN F.

PA (CONL/) CONLEY P B.

PA (TOPP/) TOPPER J N.

PA (LAWD/) LAW D.

XX Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

XX WPI; 2002-626554/67.

DR N-PSDB; ABQ98651.

XX New polypeptide designated ORFX are present in human atherogenic cells

PT and are useful to prevent and treat ORFX-associated disorders including

PT cancer, allergy, wound healing or autoimmune, cardiovascular or

PT inflammatory disease.

XX Claim 10; SEQ ID NO 916; 78pp; English.

PS The present invention relates to novel human ORFX polypeptides and their

CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences

CC were discovered in human atherogenic cells, in particular in platelets

CC and human umbilical vein endothelial cells (HUVEC) and are expressed in

CC many other tissues as well. Atherogenic cells are cells which have the

CC potential to develop atherosclerotic plaques. The ORFX polypeptides and

CC nucleic acids are useful for treating or preventing a pathological

CC condition associated with an ORFX-associated disorder, e.g. cancer,

CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood

CC coagulation disorders or inflammatory disorders. Note: The sequence data

CC for this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from the USPTO web site at

CC seqdata.uspto.gov/sequence.html?docid=2002082206

XX Sequence 55 AA;

SQ

Query Match 75.6%; Score 34; DB 5; Length 55;

Best Local Similarity 77.8%; Pred. No. 20;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLSTYDLFV 9  
|:|:|:|:|  
Db 16 FLSTDLFVL 24

RESULT 32

AAQ73909

ID AAG73909 standard; protein; 98 AA.  
XX  
XX AAG73909;  
AC  
XX 03-SEP-2001 (first entry)  
DT  
XX Human colon cancer antigen protein SEQ ID NO:4673.  
XX  
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KM colorectal carcinoma.  
XX  
XX Homo sapiens.  
OS  
XX MO200122920-A2.  
FN  
XX 05-APR-2001.  
PD  
XX 28-SEP-2000; 2000WO-US026524.  
PF  
XX 29-SEP-1999; 99US-0157137P.  
PR 03-NOV-1999; 99US-0163280P.  
XX  
XX (HUMA-) HUMAN GENOME SCL INC.  
PA  
XX Ruben SM, Barash SC, Birse CE, Rosen CA;  
FI  
XX MPI; 2001-235357/24.  
DR N-PSDB; AAH33340.  
DR  
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers.  
XX  
XX Claim 11; Page 6473-6474; 9803jp; English.  
PS  
XX AAH32963 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the  
CC proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene therapy  
CC and vaccine production. N and P may be used in the prevention, diagnosis  
CC and treatment of diseases associated with inappropriate P expression. For  
CC example, N and P may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of P by expressing inactive proteins or to  
CC supplement the patient's own production of P. Additionally, N may be used  
CC to produce the colon cancer-associated P, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the proteins. N and P  
CC can be used in the prevention, diagnosis and treatment of colorectal  
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent  
CC sequences used in the exemplification of the present invention. N.B.  
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at  
CC time of publication, meaning no sequences are present for SEQ ID NO:1027  
CC to 1052, 7921 and 7922  
XX  
SQ Sequence 98 AA;  
Query Match 75.6%; Score 34; DB 4; Length 98;  
Best Local Similarity 77.8%; Pred. No. 37;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
| | | | |  
| | | | |  
Db 16 FLSIDLFLV 24

RESULT 33  
ADC32867  
ID ADC32867 standard; protein; 160 AA.  
XX  
XX ADC32867;  
AC  
XX 18-DEC-2003 (first entry)  
DT  
XX Human novel contig-encoded polypeptide sequence, SEQ ID NO:2949.  
DE

XX  
KM Human; diagnostic; drug screening; forensics; gene mapping;  
KM biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
KM neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
KM ulcers; osteoporosis; autoimmune disease; cancer;  
KM molecular weight marker; food supplement; antiparkinsonian; noctropic;  
KM neuroprotective; antianaemic; anticoagulant; thrombolytic; vlnarary;  
KM antilicer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;  
KM gene therapy; chromosome X.  
XX  
XX Homo sapiens.  
OS  
XX MO2003029271-A2.  
FN  
XX 10-APR-2003.  
PD  
XX 24-SEP-2002; 2002WO-US030474.  
PF  
XX 24-SEP-2001; 2001US-0324631P.  
PR  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;  
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
PI Haley-Vicente D, Dymnac RT;  
XX  
XX MPI; 2003-371981/35.  
DR N-PSDB; ADC32100.  
DR  
XX New polynucleotide and polypeptide useful for diagnosing, preventing or  
PT treating conditions such as neurodegenerative diseases, anaemia, platelet  
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
PT cancer.  
XX  
XX Example 2; SEQ ID NO 2949; 1185bp; English.  
PS  
XX The invention relates to 971 novel human cDNA sequences (ADC29919-  
XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
CC invention also relates to nucleic acid sequences over 99% identical with  
CC the novel human cDNAs. The invention additionally encompasses expression  
CC vectors and host cells comprising a nucleic acid of the invention; the  
CC recombinant production of a polypeptide of the invention; an antibody  
CC against a polypeptide of the invention; a method of detecting  
CC polynucleotides or polypeptides of the invention; and methods of  
CC identifying a compound which binds to a polypeptide of the invention. The  
CC invention further discloses methods of preventing, treating or  
CC ameliorating a medical condition; kits comprising polynucleotide probes  
CC and/or monoclonal antibodies for carrying out the methods of the  
CC invention; methods for the identification of compounds that modulate the  
CC expression or activity of the polynucleotide and/or polypeptide; and 767  
CC contig sequences corresponding to the cDNA sequences of the invention.  
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
CC -ADC33594). The nucleic acids and polypeptides of the invention are  
CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
CC identification of mutations responsible for genetic disorders or other  
CC traits, for assessing biodiversity, and in producing many other types of  
CC data and products dependent on DNA and amino acid sequences. They are  
CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
CC disease and other neurodegenerative diseases, anaemia, platelet  
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
CC cancer. The nucleic acids may also be used as hybridisation probes or  
CC primers, and in the recombinant production of a protein. The polypeptides  
CC are also useful in generating antibodies, as molecular weight markers,  
CC and as food supplements. The present sequence represents a human contig-  
CC encoded polypeptide sequence used in an example of the invention. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 160 AA;  
Query Match 75.6%; Score 34; DB 7; Length 160;  
Best Local Similarity 77.8%; Pred. No. 62;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLSYDLFV 9  
 DB 78 FLSYDLFVL 86

RESULT 34  
 ADM04330  
 ID ADM04330 standard; protein; 247 AA.

XX ADM04330;

DT 20-MAY-2004 (first entry)

DE Human protein of the invention SEQ ID NO:3015.

XX human; gene therapy; diagnostic marker; pharmaceutical.

OS Homo sapiens.

PN EP1347046-A1.

XX 24-SEP-2003.

XX 12-APR-2002; 2002EP-00008400.

XX 22-MAR-2002; 2002JP-00137785.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;

XX WPI; 2003-723558/69.

DR N-PSDB; ADM01887.

XX New polynucleotides and polypeptides are useful in gene therapy, for  
 PT developing a diagnostic marker or medicines for regulating their  
 PT expression and activity, or as a target of gene therapy.

PT Claim 1; SEQ ID NO 3015; 305pp; English.

XX The invention relates to a novel human polynucleotide and the encoded  
 CC polypeptide. A polynucleotide of the invention may have a use in gene  
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
 CC as a primer for synthesizing the polynucleotide or as a probe for  
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
 CC useful in gene therapy, for developing a diagnostic marker or medicines  
 CC for regulating their expression and activity, or as a target of gene  
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
 CC are useful as pharmaceutical agents. The present sequence represents a  
 CC protein sequence of the invention.

XX Sequence 247 AA;

XX Query Match

Best Local Similarity 66.7%; Score 34; DB 7; Length 247;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9

DB 47 FLSYDLFV 55

XX RESULT 35

ADA34406

XX ADA34406 standard; protein; 363 AA.

AC ADA34406;

XX 20-NOV-2003 (first entry)

XX

XX

XX

XX

XX

XX Acinetobacter baumannii protein #1567.

DE Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;  
 XX plant biocontrol agent.

XX Acinetobacter baumannii.

OS US6562958-B1.

PN 13-MAY-2003.

XX 04-JUN-1999; 99US-00328352.

XX 09-JUN-1998; 98US-0088701P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton G, Bush D;

XX WPI; 2003-576092/54.

DR N-PSDB; ADA30280.

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
 PT for diagnosing a bacterial disease, as components of antibacterial  
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
 PT plants.

XX Example; SEQ ID NO 5693; 328pp; English.

XX The invention relates to isolated Acinetobacter baumannii nucleic acids.  
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents  
 CC for diagnosing a bacterial disease, as components of antibacterial  
 CC vaccines, as targets for antibacterial drugs, to detect the presence of  
 CC A. baumannii and other Acinetobacter species in a sample, in screening  
 CC compounds for the ability to interfere with the A. baumannii life cycle  
 CC or to inhibit A. baumannii infection, and as biocontrol agents for  
 CC plants. The present sequence represents the amino acid sequence of an A.  
 CC baumannii protein.

XX Sequence 363 AA;

XX Query Match

Best Local Similarity 75.6%; Score 34; DB 6; Length 363;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9

DB 39 FLSYDLFV 47

XX RESULT 36

AAW98765

XX AAW98765 standard; protein; 412 AA.

AC AAW98765;

XX 31-MAR-1999 (first entry)

XX H. pylori GHPD 1113 protein.

XX GHPD protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 KW peptic ulcer disease.

XX Helicobacter pylori.

OS WO9843478-A1.

PN 08-OCT-1998.

XX 01-APR-1998; 98WO-US006371.

XX 01-APR-1997; 97US-00833457.

XX

XX

XX

XX

XX

PR 24-JUN-1997; 97US-00881227.  
PR 29-JUL-1997; 97US-00902615.  
XX (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Kleanthouse H, Al-Garawi A, Miller C, Tomb J, Oomen RP,  
XX  
DR WPI; 1998-542293/46.  
DR N-PSDB; AAX14484.  
XX  
PT New isolated Helicobacter polynucleotides - used to develop products for  
PT the diagnosis, prevention and treatment of Helicobacter infections and  
PT gastrointestinal diseases.  
XX  
PS Claim 8; Page 1651-1653; 2054pp; English.  
XX  
CC This sequence represents a Helicobacter pylori GHPD protein of the  
CC invention. The polypeptides can be used for preventing or treating  
CC Helicobacter infections, and gastroduodenal diseases associated with  
CC these infections, including acute, chronic, and atrophic gastritis, and  
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
CC used for the production of antibodies. The products can also be used for  
CC detection and diagnosis  
XX  
SQ Sequence 412 AA;  
  
Query Match 75.6%; Score 34; DB 2; Length 412;  
Best Local Similarity 77.8%; Pred. No. 1.7e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 FLSTYDLFPV 9  
Db |||:||||  
253 FLSTKDFV 261  
  
RESULT 37  
ABB93691  
ID ABB93691 standard; protein; 516 AA.  
XX  
AC ABB93691;  
XX  
XX 31-MAY-2002 (first entry)  
XX  
DE Herbicidally active polypeptide SEQ ID NO 2902.  
XX  
KM Herbicidal; plant; agriculture; herbicide.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200210210-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 28-AUG-2001; 2001WO-EP009892.  
XX  
PR 28-AUG-2001; 2001WO-EP009892.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Tietjen K, Weidner M;  
XX  
DR WPI; 2002-269010/31.  
XX  
PT Identifying plant target proteins for herbicidally active compounds,  
PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
PT organisms.  
XX  
PS Claim 5; SEQ ID NO 2902; 261pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
CC for herbicidally active compounds, comprising aligning and comparing

CC nucleic acid or amino acid sequences from plant with nucleic acid or  
CC amino acid sequences from non-plant organisms using suitable search  
CC parameters, where plant sequences having an E-value greater by a factor  
CC of 3 than the E-value of most similar non-plant sequences are selected.  
CC The polypeptides or nucleic acids encoding them are useful for  
CC identifying modulators. The identified modulators are useful as  
CC herbicides  
XX  
SQ Sequence 516 AA;  
  
Query Match 75.6%; Score 34; DB 5; Length 516;  
Best Local Similarity 85.7%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 SYDLFPV 9  
Db |||:||||  
37 SYDLFVI 43  
  
RESULT 38  
ADM05144  
ID ADM05144 standard; protein; 576 AA.  
XX  
AC ADM05144;  
XX  
XX 20-MAY-2004 (first entry)  
XX  
DE Human protein of the invention SEQ ID NO:3829.  
XX  
KM human; gene therapy; diagnostic marker; pharmaceutical.  
XX  
OS Homo sapiens.  
XX  
PN EP1347046-A1.  
XX  
PD 24-SEP-2003.  
XX  
PF 12-APR-2002; 2002EP-00008400.  
XX  
PR 22-MAR-2002; 2002JP-00137785.  
XX  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,  
XX  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,  
XX  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;  
XX  
DR WPI; 2003-723558/69.  
DR N-PSDB; ADM02701.  
XX  
PT New polynucleotides and polypeptides are useful in gene therapy, for  
PT developing a diagnostic marker or medicines for regulating their  
PT expression and activity, or as a target of gene therapy.  
XX  
PS Claim 1; SEQ ID NO 3829; 305pp; English.  
XX  
CC The invention relates to a novel human polynucleotide and the encoded  
CC polypeptide. A polynucleotide of the invention may have a use in gene  
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
CC as a primer for synthesizing the polynucleotide or as a probe for  
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
CC useful in gene therapy, for developing a diagnostic marker or medicines  
CC for regulating their expression and activity, or as a target of gene  
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides  
CC are useful as pharmaceutical agents. The present sequence represents a  
CC protein sequence of the invention.  
XX  
SQ Sequence 576 AA;  
  
Query Match 75.6%; Score 34; DB 7; Length 576;  
Best Local Similarity 66.7%; Pred. No. 2.5e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTDLPV 9  
 DB 376 FLSTDLPV 384

RESULT 39  
 ADA55118  
 ID ADA55118 standard; protein; 674 AA.  
 AC ADA55118;  
 XX 20-NOV-2003 (first entry)  
 DE Human protein, SEQ ID 2666.  
 XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
 KW Gene therapy; human; secretory protein; membrane proteins; cancer;  
 KM inflammatory disease; osteoporosis; neurological disease.  
 XX Homo sapiens.  
 OS  
 XX EPI293569-A2.  
 XX 19-MAR-2003.  
 XX 21-MAR-2002; 2002EP-00006586.  
 XX 14-SEP-2001; 2001JP-00328381.  
 PR 24-JAN-2002; 2002US-0350435P.  
 XX (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;  
 XX WPI: 2003-395539/38.  
 DR N-PSDB; ADA53479.  
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.  
 XX Claim 14; SEQ ID NO 2686; 205pp; English.  
 CC The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.  
 XX SQ Sequence 674 AA;

QY 1 FLSTDLPV 9  
 DB 474 FLSTDLPV 482

Query Match 75.6%; Score 34; DB 6; Length 674;  
 Best Local Similarity 66.7%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 40  
 ADM72293  
 ID ADM72293 standard; protein; 1031 AA.  
 AC ADM72293;  
 XX 17-JUN-2004 (first entry)  
 DE Equine TLR9 polypeptide.

KW TLR9; toll-like receptor 9; CpG; TLR9 signaling; equine.  
 OS Equus caballus.  
 XX MO2004026888-A2.  
 XX 01-APR-2004.  
 PD 19-SEP-2003; 2003WO-US029577.  
 PF 19-SEP-2002; 2002US-0412479P.  
 XX (COLE-) COLEY PHARM GMBH.  
 PA (UYSA-) UNIV SASKATCHEWAN.  
 PA (QIAG-) QIAGEN GMBH.  
 XX Lipford GB, Mookerjee N, Babiuk L, Brownlie R, Griebel P;  
 PI Mutwiri G, Hecker R;  
 XX WPI: 2004-295374/27.  
 DR N-PSDB; ADM72295.  
 PT New polypeptide, useful for identifying key amino acids in a TLR9 of a  
 PT first species which confer specificity for CpG DNA optimized for TLR9 of  
 PT the first species.  
 XX Claim 1; SEQ ID NO 13; 170pp; English.

CC The invention relates to novel Toll-like receptor 9 (TLR9) polypeptides  
 CC and encoding polynucleotides. It provides methods for (i) identifying key  
 CC amino acids in a TLR9 of a first species which confer specificity for CpG  
 CC DNA optimized for TLR9 of the first species; (ii) identifying key amino  
 CC acids in human TLR9 which confer specificity for CpG DNA optimized for  
 CC human TLR9; (iii) identify a TLR9 ligand; and (iv) identify species-  
 CC specific CpG-motif preference of the isolated polypeptide. The screening  
 CC method to identify species-specific CpG-motif preference of the isolated  
 CC polypeptide comprises: contacting an isolated polypeptide with a CpG DNA  
 CC comprising a hexamer sequence consisting of GACGTT, AGCGTT, CACGTT,  
 CC TACGTT, GCGGTT, GCGGTT, GTCGTT, GATGTT, GATGTT, GACGTT, GACGTT,  
 CC GACGTT, GACGCT, GACGAT, GACGCT, GACGCT, GACGCT, GACGCT, GACGCT,  
 CC signal in response to the contacting; and identifying a species-specific  
 CC CpG-motif preference when the signal in response to the contacting is  
 CC consistent with TLR9 signaling. The signal comprises expression of a  
 CC reporter gene responsive to TLR/IL-1R signal transduction pathway. The  
 CC reporter gene is operatively linked to a promoter sensitive to NF- $\kappa$ B. The  
 CC CpG DNA is an oligodeoxynucleotide having a sequence consisting of  
 CC sequences selected from ADM72319-ADM72317. The polypeptide is useful for  
 CC identifying key amino acids in a TLR9 of a first species which confer  
 CC specificity for CpG DNA optimized for TLR9 of the first species. The  
 CC present sequence represents an equine TLR9 polypeptide.  
 XX SQ Sequence 1031 AA;

QY 2 LSYDLPV 9  
 DB 866 LSYDLPV 873

Query Match 75.6%; Score 34; DB 8; Length 1031;  
 Best Local Similarity 87.5%; Pred. No. 4.6e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Search completed: January 12, 2005, 20:07:04  
 Job time : 212.6 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2005, 20:16:00 ; Search time 93.6 Seconds  
(without alignments)  
34.739 Million cell updates/sec

Title: US-09-870-216C-9  
Perfect score: 45  
Sequence: 1 FLSYDLFV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubppaa/PCR\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubppaa/CTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	9	11 US-09-870-216C-9	Sequence 9, Appl1
2	45	100.0	9	13 US-10-017-327-9	Sequence 9, Appl1
3	36	80.0	51	13 US-09-833-245-2059	Sequence 2059, Ap
4	36	80.0	51	14 US-10-060-255-62	Sequence 62, Appl
5	36	80.0	51	17 US-10-881-088-62	Sequence 62, Appl
6	35	77.8	130	9 US-09-729-674-16	Sequence 16, Appl
7	35	77.8	130	17 US-10-913-553-16	Sequence 16, Appl
8	35	77.8	162	11 US-09-978-360A-732	Sequence 732, App
9	35	77.8	162	14 US-10-319-763-126	Sequence 126, App
10	35	77.8	162	14 US-10-319-763-126	Sequence 220, App
11	35	77.8	162	14 US-10-170-385-53	Sequence 53, Appl
12	35	77.8	162	15 US-10-264-237-1995	Sequence 1995, Ap
13	35	77.8	163	10 US-09-809-391-462	Sequence 462, App

14	35	77.8	163	10	US-09-882-171-462	Sequence 462, App
15	35	77.8	163	9	US-10-164-861-462	Sequence 462, App
16	34	75.6	55	9	US-09-867-550-916	Sequence 916, App
17	34	75.6	66	15	US-10-424-599-175697	Sequence 175697, App
18	34	75.6	98	14	US-10-106-698-4683	Sequence 4683, App
19	34	75.6	247	15	US-10-108-260A-3015	Sequence 3015, App
20	34	75.6	346	15	US-10-424-599-217530	Sequence 217530, App
21	34	75.6	412	10	US-09-882-227-396	Sequence 396, App
22	34	75.6	576	15	US-10-108-360A-3829	Sequence 3829, App
23	34	75.6	674	14	US-10-094-749-2686	Sequence 2686, App
24	33	73.3	56	15	US-10-424-599-278654	Sequence 278654, App
25	33	73.3	67	15	US-10-424-599-284635	Sequence 284635, App
26	33	73.3	88	17	US-10-425-115-318398	Sequence 318398, App
27	33	73.3	100	17	US-10-425-115-195354	Sequence 195354, App
28	33	73.3	211	14	US-10-083-357-691	Sequence 691, App
29	33	73.3	573	14	US-10-369-493-2314	Sequence 2314, App
30	33	73.3	573	14	US-10-369-493-2320	Sequence 2320, App
31	33	73.3	583	14	US-10-369-493-22678	Sequence 22678, App
32	33	73.3	839	15	US-10-424-599-206656	Sequence 206656, App
33	33	73.3	1507	15	US-10-424-599-206659	Sequence 206659, App
34	32	71.1	62	16	US-10-437-963-177564	Sequence 177564, App
35	32	71.1	80	17	US-10-425-115-256908	Sequence 256908, App
36	32	71.1	153	15	US-10-424-599-147896	Sequence 147896, App
37	32	71.1	165	14	US-10-156-761-8778	Sequence 8778, App
38	32	71.1	265	15	US-10-336-603A-52	Sequence 52, App
39	32	71.1	279	9	US-09-778-927A-64	Sequence 64, Appl
40	32	71.1	288	15	US-10-336-603A-48	Sequence 48, Appl
41	32	71.1	291	16	US-10-437-963-131686	Sequence 131686, App
42	32	71.1	305	15	US-10-336-603A-46	Sequence 46, Appl
43	32	71.1	313	17	US-10-739-930-6831	Sequence 6831, App
44	32	71.1	328	15	US-10-336-603A-44	Sequence 44, Appl
45	32	71.1	328	15	US-10-336-603A-50	Sequence 50, Appl

# ALIGNMENTS

RESULT 1  
US-09-870-216C-9  
Sequence 9, Application US/09870216C  
Publication No. US20040138135A1  
GENERAL INFORMATION:  
APPLICANT: Charles A. Nicolette  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
FILE REFERENCE: 68126881210100  
CURRENT APPLICATION NUMBER: US/09/870,216C  
CURRENT FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 60/209,391  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: 60/226,256  
PRIOR FILING DATE: 2000-08-17  
PRIOR APPLICATION NUMBER: 60/257,008  
PRIOR FILING DATE: 2000-12-20  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-870-216C-9

Query Match  
Best Local Similarity 100.0%; Score 45; DB 11; Length 9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0;

Qy 1 FLSYDLFV 9  
Db 1 FLSYDLFV 9

RESULT 2  
US-10-017-327-9  
Sequence 9, Application US/10017327

Publication No. US20020155471A1  
GENERAL INFORMATION:  
APPLICANT: Charles A. Nicolette  
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND  
FILE REFERENCE: GZ 2101.20  
CURRENT APPLICATION NUMBER: US/10/017,327  
CURRENT FILING DATE: 2001-12-06  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-017-327-9

Query Match 100.0%; Score 45; DB 13; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9  
Db 1 FLSYDLFV 9

RESULT 3  
US-09-833-245-2059  
Sequence 2059, Application US/09833245  
Publication No. US20040010134A1  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PF546PCT  
CURRENT APPLICATION NUMBER: US/09/833,245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229,358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256,931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199,384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2059  
LENGTH: 51  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-833-245-2059

Query Match 80.0%; Score 36; DB 11; Length 51;  
Best Local Similarity 66.7%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9  
Db 7 FVSIDYFIV 15

RESULT 4  
US-10-060-255-62  
Sequence 62, Application US/10060255  
Publication No. US20030113840A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 25 Human secreted proteins  
FILE REFERENCE: P2042P1  
CURRENT APPLICATION NUMBER: US/10/060,255  
CURRENT FILING DATE: 2002-02-01  
PRIOR APPLICATION NUMBER: 09/781,417  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: PCT/US00/22325  
PRIOR FILING DATE: 2000-08-16  
PRIOR APPLICATION NUMBER: 60/149,182

PRIOR FILING DATE: 1999-08-17  
NUMBER OF SEQ ID NOS: 86  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 62  
LENGTH: 51  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-060-255-62

Query Match 80.0%; Score 36; DB 14; Length 51;  
Best Local Similarity 66.7%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9  
Db 7 FVSIDYFIV 15

RESULT 5  
US-10-881-088-62  
Sequence 62, Application US/10881088  
Publication No. US20040241803A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 25 Human secreted proteins  
FILE REFERENCE: P2042P1  
CURRENT APPLICATION NUMBER: US/10/881,088  
CURRENT FILING DATE: 2004-07-01  
PRIOR APPLICATION NUMBER: US/09/781,417  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: PCT/US00/22325  
PRIOR FILING DATE: 2000-08-16  
PRIOR APPLICATION NUMBER: 60/149,182  
PRIOR FILING DATE: 1999-08-17  
NUMBER OF SEQ ID NOS: 86  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 62  
LENGTH: 51  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-881-088-62

Query Match 80.0%; Score 36; DB 17; Length 51;  
Best Local Similarity 66.7%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9  
Db 7 FVSIDYFIV 15

RESULT 6  
US-09-729-674-16  
Sequence 16, Application US/09729674  
Patent No. US2001003935A1  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: McCoy, John M.  
APPLICANT: Lavallee, Edward R.  
APPLICANT: Collins-Racie, Lisa A.  
APPLICANT: Evans, Cheryl  
APPLICANT: Meiberg, David  
APPLICANT: Treacy, Maurice  
APPLICANT: Agostino, Michael J.  
APPLICANT: Steindinger II, Robert J.  
APPLICANT: Spaulding, Vikki  
APPLICANT: Wong, Gordon G.  
APPLICANT: Clark, Hilary  
APPLICANT: Fechtel, Kim  
APPLICANT: Genetics Institute, Inc.  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
FILE REFERENCE: 6055-64X  
CURRENT APPLICATION NUMBER: US/09/729,674



;; CURRENT FILING DATE: 2000-12-04  
;; PRIOR APPLICATION NUMBER: 09/539,330  
;; PRIOR FILING DATE: 2000-03-30  
;; NUMBER OF SEQ ID NOS: 283  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 16  
;; LENGTH: 130  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-729-674-16

Query Match 77.8%; Score 35; DB 9; Length 130;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
Db 18 LSYDLFV 24

## RESULT 7

US-10-913-553-16  
;; Sequence 16, Application US/10913553  
;; Publication No. US2005003491A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Jacobs, Kenneth  
;; APPLICANT: McCoy, John M.  
;; APPLICANT: Lavallie, Edward R.  
;; APPLICANT: Collins-Racie, Lisa A.  
;; APPLICANT: Evans, Cheryl  
;; APPLICANT: Merberg, David  
;; APPLICANT: Treacy, Maurice  
;; APPLICANT: Agostino, Michael J.  
;; APPLICANT: Steindinger II, Robert J.  
;; APPLICANT: Spaulding, Vikki  
;; APPLICANT: Wong, Gordon G.  
;; APPLICANT: Clark, Hilary  
;; APPLICANT: Fechtel, Kim  
;; APPLICANT: Genetics Institute, Inc.  
;; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM  
;; FILE REFERENCE: 6055-64X  
;; CURRENT APPLICATION NUMBER: US/10/913,553  
;; CURRENT FILING DATE: 2004-08-09  
;; PRIOR APPLICATION NUMBER: US/09/539,330  
;; PRIOR FILING DATE: 2000-03-30  
;; NUMBER OF SEQ ID NOS: 283  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 16  
;; LENGTH: 130  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-913-553-16

Query Match 77.8%; Score 35; DB 17; Length 130;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
Db 18 LSYDLFV 24

## RESULT 8

US-09-978-360A-732  
;; Sequence 732, Application US/09978360A  
;; Publication No. US20040110939A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Edwards, Jean-Baptiste Dumas Milne  
;; APPLICANT: Duclet, Aymeric  
;; APPLICANT: Bougueleret, Lydie  
;; APPLICANT: Joberet, Severin  
;; APPLICANT: Clusel, Catherine  
;; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides

;; FILE REFERENCE: 56.US4.CIP  
;; CURRENT APPLICATION NUMBER: US/09/978,360A  
;; CURRENT FILING DATE: 2001-10-15  
;; PRIOR APPLICATION NUMBER: US 60/066,677  
;; PRIOR FILING DATE: 1997-11-13  
;; PRIOR APPLICATION NUMBER: US 60/069,957  
;; PRIOR FILING DATE: 1997-12-17  
;; PRIOR APPLICATION NUMBER: US 60/074,121  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: US 60/081,563  
;; PRIOR FILING DATE: 1998-04-13  
;; PRIOR APPLICATION NUMBER: US 60/096,116  
;; PRIOR FILING DATE: 1998-08-10  
;; PRIOR APPLICATION NUMBER: US 60/099,273  
;; PRIOR FILING DATE: -09-04  
;; PRIOR APPLICATION NUMBER: US 09/191,997  
;; PRIOR FILING DATE: 1998-11-13  
;; PRIOR APPLICATION NUMBER: US 09/215,435  
;; PRIOR FILING DATE: 1998-12-17  
;; PRIOR APPLICATION NUMBER: PCT/IB98/02122  
;; PRIOR FILING DATE: 1998-12-17  
;; PRIOR APPLICATION NUMBER: US 09/247,155  
;; PRIOR FILING DATE: 1999-02-09  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 810  
;; SOFTWARE: Patent .pm  
;; SEQ ID NO 732  
;; LENGTH: 162  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: SIGNAL  
;; LOCATION: -94...-1  
US-09-978-360A-732

Query Match 77.8%; Score 35; DB 11; Length 162;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
Db 18 LSYDLFV 24

## RESULT 9

US-10-319-763-126  
;; Sequence 126, Application US/10319763  
;; Publication No. US20030144490A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
;; APPLICANT: Bougueleret, Lydie  
;; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS  
;; FILE REFERENCE: G-031.US04.DIV  
;; CURRENT APPLICATION NUMBER: US/10/319,763  
;; CURRENT FILING DATE: 2002-12-10  
;; PRIOR APPLICATION NUMBER: 60/066,677  
;; PRIOR FILING DATE: 1997-11-13  
;; PRIOR APPLICATION NUMBER: 60/069,957  
;; PRIOR FILING DATE: 1997-12-17  
;; PRIOR APPLICATION NUMBER: 60/074,121  
;; PRIOR FILING DATE: 1998-02-09  
;; PRIOR APPLICATION NUMBER: 60/081,563  
;; PRIOR FILING DATE: 1998-04-13  
;; PRIOR APPLICATION NUMBER: 60/096,116  
;; PRIOR FILING DATE: 1998-08-10  
;; PRIOR APPLICATION NUMBER: 60/099,273  
;; PRIOR FILING DATE: 1998-09-04  
;; NUMBER OF SEQ ID NOS: 229  
;; SOFTWARE: Patent .pm  
;; SEQ ID NO 126  
;; LENGTH: 162  
;; TYPE: PRT

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -21...-1  
US-10-319-763-126

Query Match 77.8%; Score 35; DB 14; Length 162;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSYDLFV 8  
Db 18 LSYDLFV 24

RESULT 10  
US-10-319-763-220  
Sequence 220, Application US/10319763  
Publication No. US20030144490A1  
GENERAL INFORMATION:  
APPLICANT: Dumas Maline Edwards, Jean-Baptiste  
APPLICANT: Duclet, Aymeric  
TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS  
FILE REFERENCE: G-031.US04.DIV  
CURRENT APPLICATION NUMBER: US/10/319,763  
CURRENT FILING DATE: 2002-12-10  
PRIOR APPLICATION NUMBER: 60/066,677  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/069,957  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/074,121  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/081,563  
PRIOR FILING DATE: 1998-04-13  
PRIOR APPLICATION NUMBER: 60/096,116  
PRIOR FILING DATE: 1998-08-10  
PRIOR APPLICATION NUMBER: 60/099,273  
PRIOR FILING DATE: 1998-09-04  
NUMBER OF SEQ ID NOS: 229  
SOFTWARE: Patent.pm  
SEQ ID NO 220  
LENGTH: 162  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -94...-1  
US-10-319-763-220

Query Match 77.8%; Score 35; DB 14; Length 162;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSYDLFV 8  
Db 18 LSYDLFV 24

RESULT 11  
US-10-170-385-53  
Sequence 53, Application US/10170385  
Publication No. US20030203372A1  
GENERAL INFORMATION:  
APPLICANT: Ward, Neil Raymond  
APPLICANT: Mundy, Christopher Robert  
APPLICANT: Kan, On  
APPLICANT: Harris, Robert Alan  
APPLICANT: White, Jonathan  
APPLICANT: Binley, Katie Mary  
APPLICANT: Rayner, William Nigel  
APPLICANT: Naylor, Stuart  
APPLICANT: Kingman, Susan Mary

APPLICANT: Krige, David  
TITLE OF INVENTION: ANALYSIS METHOD  
FILE REFERENCE: 532682000100  
CURRENT APPLICATION NUMBER: US/10/170,385  
CURRENT FILING DATE: 2002-06-12  
PRIOR APPLICATION NUMBER: PCT/GB02/01662  
PRIOR FILING DATE: 2002-04-08  
PRIOR APPLICATION NUMBER: PCT/GB01/05458  
PRIOR FILING DATE: 2001-12-10  
NUMBER OF SEQ ID NOS: 549  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 53  
LENGTH: 162  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-170-385-53

Query Match 77.8%; Score 35; DB 14; Length 162;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSYDLFV 8  
Db 18 LSYDLFV 24

RESULT 12  
US-10-264-237-1995  
Sequence 1995, Application US/10264237  
Publication No. US20040009491A1  
GENERAL INFORMATION:  
APPLICANT: Birse et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA131PI  
CURRENT APPLICATION NUMBER: US/10/264,237  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/16450  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205,515  
PRIOR FILING DATE: 2000-05-19  
NUMBER OF SEQ ID NOS: 2876  
SOFTWARE: Patentln Ver. 3.1  
SEQ ID NO 1995  
LENGTH: 162  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-264-237-1995

Query Match 77.8%; Score 35; DB 15; Length 162;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSYDLFV 8  
Db 18 LSYDLFV 24

RESULT 13  
US-09-809-391-462  
Sequence 462, Application US/09809391  
Publication No. US20030049618A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P2  
CURRENT APPLICATION NUMBER: US/09/809,391  
CURRENT FILING DATE: 2001-03-16  
Prior application data removed - consult PAM or file wrapper  
NUMBER OF SEQ ID NOS: 761  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 462  
LENGTH: 163  
TYPE: PRT

ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: (163)  
 OTHER INFORMATION: Xaa equals stop translation  
 US-09-809-391-462

Query Match 77.8%; Score 35; DB 10; Length 163;  
 Best Local Similarity 100.0%; Pred. No. 92;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSYDLFV 8  
 |||||  
 Db 18 LSYDLFV 24

RESULT 14  
 US-09-882-171-462  
 ; Sequence 462, Application US/09882171  
 ; Publication No. US20030175858A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruben et al.  
 ; TITLE OF INVENTION: 186 Human Secreted proteins  
 ; FILE REFERENCE: P20022  
 ; CURRENT APPLICATION NUMBER: US/09/882,171  
 ; CURRENT FILING DATE: 2001-06-18  
 ; PRIOR APPLICATION NUMBER: 09/809,391  
 ; PRIOR FILING DATE: 2001-03-16  
 ; PRIOR APPLICATION NUMBER: 09/149,476  
 ; PRIOR FILING DATE: 1998-09-08  
 ; PRIOR APPLICATION NUMBER: PCT/US98/04493  
 ; PRIOR FILING DATE: 1998-03-06  
 ; PRIOR APPLICATION NUMBER: 60/040,162  
 ; PRIOR FILING DATE: 1997-03-07  
 ; PRIOR APPLICATION NUMBER: 60/040,333  
 ; PRIOR FILING DATE: 1997-03-07  
 ; PRIOR APPLICATION NUMBER: 60/038,621  
 ; PRIOR FILING DATE: 1997-03-07  
 ; PRIOR APPLICATION NUMBER: 60/040,626  
 ; PRIOR FILING DATE: 1997-03-07  
 ; PRIOR APPLICATION NUMBER: 60/040,334  
 ; PRIOR FILING DATE: 1997-03-07  
 ; PRIOR APPLICATION NUMBER: 60/040,336  
 ; PRIOR FILING DATE: 1997-03-07  
 ; PRIOR APPLICATION NUMBER: 60/040,163  
 ; PRIOR FILING DATE: 1997-03-07  
 ; PRIOR APPLICATION NUMBER: 60/047,600  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,615  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,597  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,502  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,633  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,583  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,617  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,618  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,503  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,592  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,581  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,584  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,500  
 ; PRIOR FILING DATE: 1997-05-23  
 ; PRIOR APPLICATION NUMBER: 60/047,587

;; PRIOR FILING DATE: 1997-05-23  
 ;; PRIOR APPLICATION NUMBER: 60/047,492  
 ;; PRIOR FILING DATE: 1997-05-23  
 ;; PRIOR APPLICATION NUMBER: 60/047,598  
 ;; PRIOR FILING DATE: 1997-05-23  
 ;; PRIOR APPLICATION NUMBER: 60/047,613  
 ;; PRIOR FILING DATE: 1997-05-23  
 ;; PRIOR APPLICATION NUMBER: 60/047,582  
 ;; PRIOR FILING DATE: 1997-05-23  
 ;; PRIOR APPLICATION NUMBER: 60/047,596  
 ;; PRIOR FILING DATE: 1997-05-23  
 ;; PRIOR APPLICATION NUMBER: 60/047,612  
 ;; PRIOR FILING DATE: 1997-05-23  
 ;; PRIOR APPLICATION NUMBER: 60/047,632  
 ;; PRIOR FILING DATE: 1997-05-23  
 ;; PRIOR APPLICATION NUMBER: 60/047,601  
 ;; PRIOR FILING DATE: 1997-05-23  
 ;; PRIOR APPLICATION NUMBER: 60/043,580  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,568  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,314  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,569  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,311  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,671  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,674  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,669  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,312  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,313  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,672  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/043,315  
 ;; PRIOR FILING DATE: 1997-04-11  
 ;; PRIOR APPLICATION NUMBER: 60/048,974  
 ;; PRIOR FILING DATE: 1997-06-06  
 ;; PRIOR APPLICATION NUMBER: 60/056,886  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,877  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,889  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,893  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,630  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,878  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,662  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,872  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,882  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,637  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,903  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,888  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,879  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,880  
 ;; PRIOR FILING DATE: 1997-08-22  
 ;; PRIOR APPLICATION NUMBER: 60/056,894

PRIOR APPLICATION NUMBER: 60/056,911  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,636  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,874  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,910  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,864  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,631  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,845  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,892  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/057,761  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/047,595  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,599  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,588  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,585  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,586  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,590  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,594  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,589  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,593  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,614  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/043,578  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/043,576  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/047,501  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/043,670  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/056,632  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,664  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,876  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,881  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,909  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,875  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,862  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,887  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,908  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/048,964  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/057,650  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: 60/056,884  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/057,669  
PRIOR FILING DATE: 1997-09-05

Query Match 77.8%; Score 35; DB 10; Length 163;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
Db 18 LSYDLFV 24

RESULT 15  
US-10-164-861-462  
Sequence 462, Application US/10164861  
Publication No. US2003025248A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/10/164,861  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: US/09/149,476  
PRIOR FILING DATE: 1998-09-08  
PRIOR APPLICATION NUMBER: PCT/US98/04493  
PRIOR FILING DATE: 1998-03-06  
NUMBER OF SEQ ID NOS: 757  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 462  
LENGTH: 163  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (163)  
OTHER INFORMATION: Xaa equals stop translation  
US-10-164-861-462

Query Match 77.8%; Score 35; DB 14; Length 163;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
Db 18 LSYDLFV 24

RESULT 16  
US-09-867-550-916  
Sequence 916, Application US/09867550  
Patent No. US20020082206A1  
GENERAL INFORMATION:  
APPLICANT: Leach, Martin D.  
APPLICANT: Mehriban, Fuad  
APPLICANT: Conley, Pamela  
APPLICANT: Law, Debbie  
APPLICANT: Topper, James  
TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and  
FILE REFERENCE: 21402-013 (Cura-313)  
CURRENT APPLICATION NUMBER: US/09/867,550  
CURRENT FILING DATE: 2001-09-20  
PRIOR APPLICATION NUMBER: USSN 60/208,427  
PRIOR FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 2125  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 916  
LENGTH: 55  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-867-550-916

Query Match 75.6%; Score 34; DB 9; Length 55;  
Best Local Similarity 77.8%; Pred. No. 48;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTDLPV 9  
|||  
Db 16 FLSTDLPV 24

## RESULT 17

US-10-424-599-175697  
; Sequence 175697, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 175697  
; LENGTH: 66  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_129672C.1.pap  
US-10-424-599-175697

Query Match 75.6%; Score 34; DB 15; Length 66;  
Best Local Similarity 66.7%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTDLPV 9  
|||  
Db 30 FLSTDLPV 38

## RESULT 18

US-10-106-698-4683  
; Sequence 4683, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 4683  
; LENGTH: 98  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-106-698-4683

Query Match 75.6%; Score 34; DB 14; Length 98;  
Best Local Similarity 77.8%; Pred. No. 86;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTDLPV 9  
|||  
Db 16 FLSTDLPV 24

## RESULT 19

US-10-108-260A-3015  
; Sequence 3015, Application US/10108260A

; Publication No. US2004005560A1

; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US2004005560A1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3015  
; LENGTH: 247  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-3015

Query Match 75.6%; Score 34; DB 15; Length 247;  
Best Local Similarity 66.7%; Pred. No. 2.1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTDLPV 9  
|||  
Db 47 FLSTDLPV 55

## RESULT 20

US-10-424-599-217530  
; Sequence 217530, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 217530  
; LENGTH: 346  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_38457C.1.pap  
US-10-424-599-217530

Query Match 75.6%; Score 34; DB 15; Length 346;  
Best Local Similarity 77.8%; Pred. No. 3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTDLPV 9  
|||  
Db 293 FLSTDLPV 301

## RESULT 21

US-09-882-227-396  
; Sequence 396, Application US/09882227  
; Publication No. US20030158396A1  
; GENERAL INFORMATION:  
; APPLICANT: Kleantous Harold  
; APPLICANT: Al-Garawi Amal  
; APPLICANT: Miller, Charles  
; APPLICANT: Oomen, Jean-Francois  
; TITLE OF INVENTION: Identification of Polynucleotides  
; TITLE OF INVENTION: Encoding No. US20030158396A1 Helicobacter Polypeptides in the  
; FILE REFERENCE: 06132/047002  
; CURRENT APPLICATION NUMBER: US/09/882,227  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 08/902,615

PRIOR FILING DATE: 1997-07-29  
NUMBER OF SEQ ID NOS: 638  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 396  
LENGTH: 412  
TYPE: PRT  
ORGANISM: Helicobacter pylori  
US-09-882-227-396

Query Match 75.6%; Score 34; DB 10; Length 412;  
Best Local Similarity 77.8%; Pred. No. 3.5e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSTDLEFV 9  
Db 253 FLSTDLEFV 261

RESULT 22  
US-10-108-260A-3829  
Sequence 3829, Application US/10108260A  
Publication No. US20040005560A1  
GENERAL INFORMATION:  
APPLICANT: HELIX RESEARCH INSTITUTE  
TITLE OF INVENTION: No. US20040005560A1 full length cDNA  
FILE REFERENCE: H1-A0106  
CURRENT APPLICATION NUMBER: US/10/108,260A  
CURRENT FILING DATE: 2002-03-27  
NUMBER OF SEQ ID NOS: 5458  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3829  
LENGTH: 576  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-108-260A-3829

Query Match 75.6%; Score 34; DB 15; Length 576;  
Best Local Similarity 66.7%; Pred. No. 4.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSTDLEFV 9  
Db 376 FLSTDLEFV 384

RESULT 23  
US-10-094-749-2686  
Sequence 2686, Application US/10094749  
Publication No. US20030219741A1  
GENERAL INFORMATION:  
APPLICANT: ISOGAI, TAKAO  
APPLICANT: SUGIYAMA, TOMOYASU  
APPLICANT: OTSUKI, TETSUJI  
APPLICANT: WAKAMATSU, AI  
APPLICANT: SATO, HIROYUKI  
APPLICANT: ISHII, SHIZUKO  
APPLICANT: YAMAMOTO, JUN-ICHI  
APPLICANT: ISONO, YUUKO  
APPLICANT: HIO, YURI  
APPLICANT: OTSUKA, KAORU  
APPLICANT: NAGAI, KEIICHI  
APPLICANT: IRIE, RYOTARO  
APPLICANT: TAMECHIKA, ICHIRO  
APPLICANT: SEKI, NAOHITO  
APPLICANT: YOSHIKAWA, TSUTOMU  
APPLICANT: OTSUKA, MOTOTUKI  
APPLICANT: NAGAHARI, KENJI  
APPLICANT: MASUHO, YASUHIKO  
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
FILE REFERENCE: 084335/0160  
CURRENT APPLICATION NUMBER: US/10/094,749  
CURRENT FILING DATE: 2002-03-12  
PRIOR APPLICATION NUMBER: 60/350,435

PRIOR FILING DATE: 2002-01-24  
PRIOR APPLICATION NUMBER: JP 2001-328381  
PRIOR FILING DATE: 2001-09-14  
NUMBER OF SEQ ID NOS: 3381  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2686  
LENGTH: 674  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-094-749-2686

Query Match 75.6%; Score 34; DB 14; Length 674;  
Best Local Similarity 66.7%; Pred. No. 5.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSTDLEFV 9  
Db 474 FLSTDLEFV 482

RESULT 24  
US-10-424-599-278654  
Sequence 278654, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 278654  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)-(56)  
OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_93647C.1.pep  
US-10-424-599-278654

Query Match 73.3%; Score 33; DB 15; Length 56;  
Best Local Similarity 77.8%; Pred. No. 76;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSTDLEFV 9  
Db 4 FLSTDLEFV 12

RESULT 25  
US-10-424-599-284635  
Sequence 284635, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 284635  
LENGTH: 67

;; TYPE: PRT  
;; ORGANISM: Glycine max  
;; FEATURE:  
;; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_99050C.1.pep  
US-10-424-599-284635

Query Match 73.3%; Score 33; DB 15; Length 67;  
Best Local Similarity 77.8%; Pred. No. 90;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLSTDLPV 9  
DB 15 FLSTELSV 23

RESULT 26  
US-10-425-115-318398  
; Sequence 318398, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 318398  
; LENGTH: 88  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_53454C.1.pep  
US-10-425-115-318398

Query Match 73.3%; Score 33; DB 17; Length 88;  
Best Local Similarity 44.4%; Pred. No. 1.2e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLSTDLPV 9  
DB 37 FIAYDTFII 45

RESULT 27  
US-10-425-115-195354  
; Sequence 195354, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 195354  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_109751C.1.pep  
US-10-425-115-195354

Query Match 73.3%; Score 33; DB 17; Length 100;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLSTDLP 7  
DB 15 FLQYDLF 21

RESULT 28  
US-10-083-357-691  
; Sequence 691, Application US/10083357  
; Publication No. US20030054370A1  
; GENERAL INFORMATION:  
; APPLICANT: Qiangong, Zeng et al.  
; TITLE OF INVENTION: Systemic Discovery of New Genes  
; FILE REFERENCE: 032796-090  
; CURRENT APPLICATION NUMBER: US/10/083,357  
; CURRENT FILING DATE: 2002-02-27  
; NUMBER OF SEQ ID NOS: 1346  
; SEQ ID NO 691  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-10-083-357-691

Query Match 73.3%; Score 33; DB 14; Length 211;  
Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSYDLPV 9  
DB 20 LSYDLPFL 27

RESULT 29  
US-10-369-493-2314  
; Sequence 2314, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianteng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2314  
; LENGTH: 573  
; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
US-10-369-493-2314

Query Match 73.3%; Score 33; DB 14; Length 573;  
Best Local Similarity 71.4%; Pred. No. 7.5e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 SYDLPV 9  
DB 59 SYDLPFI 65

RESULT 30  
US-10-369-493-2320  
; Sequence 2320, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.

```

; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2320
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2320

Query Match
Best Local Similarity 73.3%; Score 33; DB 14; Length 573;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 SYDLFVV 9
DB 59 SYDLFII 65

RESULT 31
US-10-369-493-22678
; Sequence 22678; Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22678
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-22678

Query Match
Best Local Similarity 73.3%; Score 33; DB 14; Length 583;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 SYDLFVV 9
DB 53 SYDLFII 59

RESULT 32
US-10-424-599-206656
; Sequence 206656; Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovall, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

```

```

; SEQ ID NO 206656
; LENGTH: 839
; TYPE: PRT
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28638C.1.pep
US-10-424-599-206656

Query Match
Best Local Similarity 73.3%; Score 33; DB 15; Length 839;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSDLFVV 9
DB 286 YLDYDMFVL 294

RESULT 33
US-10-424-599-206659
; Sequence 206659; Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovall, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206659
; LENGTH: 1507
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28640C.1.pep
US-10-424-599-206659

Query Match
Best Local Similarity 73.3%; Score 33; DB 15; Length 1507;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSDLFVV 9
DB 824 YLDYDMFVL 832

RESULT 34
US-10-437-963-177564
; Sequence 177564; Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovall, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 177564
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:

```



OTHER INFORMATION: Clone ID: PAT\_MRT4530\_75202C.1.pep  
US-10-437-963-177564

Query Match 71.1%; Score 32; DB 16; Length 62;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
Db 40 LSYDLFV 46

RESULT 35  
US-10-425-115-256908  
Sequence 256908, Application US/10425115  
Publication No. US20040214272A1

GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425.115

PRIOR FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 256908

LENGTH: 80

TYPE: PRT

ORGANISM: Zea mays

FEATURE:  
NAME/KEY: unsure

LOCATION: (1) ..(80)

OTHER INFORMATION: unsure at all Xaa locations

OTHER INFORMATION: Clone ID: MRT4577\_165889C.1.pep

US-10-425-115-256908

Query Match 71.1%; Score 32; DB 17; Length 80;  
Best Local Similarity 55.6%; Pred. No. 1.7e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
Db 70 YMSYHLFV 78

RESULT 36  
US-10-424-599-147896  
Sequence 147896, Application US/10424599  
Publication No. US20040031072A1

GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424.599

PRIOR FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 147896

LENGTH: 153

TYPE: PRT

ORGANISM: Glycine max

FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_104572C.1.pep

US-10-424-599-147896

Query Match 71.1%; Score 32; DB 15; Length 153;  
Best Local Similarity 85.7%; Pred. No. 3.1e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLF 7  
Db 47 FLSYDLF 53

RESULT 37  
US-10-156-761-8778  
Sequence 8778, Application US/10156761  
Publication No. US20030119018A1

GENERAL INFORMATION:  
APPLICANT: OMURA, SATOSHI  
APPLICANT: IKEDA, HARUO  
APPLICANT: ISHIKAWA, JUN  
APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156.761

PRIOR FILING DATE: 2002-05-28

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 8778

LENGTH: 165

TYPE: PRT

ORGANISM: Streptomyces avermitilis

US-10-156-761-8778

Query Match 71.1%; Score 32; DB 14; Length 165;  
Best Local Similarity 85.7%; Pred. No. 3.4e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLF 7  
Db 46 FLSYDLF 52

RESULT 38  
US-10-336-603A-52  
Sequence 52, Application US/10336603A  
Publication No. US2004007297A1

GENERAL INFORMATION:  
APPLICANT: Alsobrook et al.  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

FILE REFERENCE: 21402-533A

CURRENT APPLICATION NUMBER: US/10/336.603A

PRIOR FILING DATE: 2003-01-03

PRIOR APPLICATION NUMBER: 09/746,491

PRIOR FILING DATE: 2000-12-20

PRIOR APPLICATION NUMBER: 10/055,569

PRIOR FILING DATE: 2001-10-26

NUMBER OF SEQ ID NOS: 169

SOFTWARE: Carbaseqlist version 0.1

SEQ ID NO 52

LENGTH: 265

TYPE: PRT

ORGANISM: Homo sapiens

US-10-336-603A-52

Query Match 71.1%; Score 32; DB 15; Length 265;  
Best Local Similarity 55.6%; Pred. No. 5.4e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
Db 219 YMSYHLFV 227

## RESULT 39

US-09-778-927A-64  
; Sequence 64, Application US/09778927A  
; Patent No. US20020068342A1  
; GENERAL INFORMATION:  
; APPLICANT: KHOSRAVI, Rami et al.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL  
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING  
; FILE REFERENCE: 2786-0160P  
; CURRENT APPLICATION NUMBER: US/09/778, 927A  
; CURRENT FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: IL 134453  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: IL135341  
; PRIOR FILING DATE: 2000-03-29  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 64  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(279 )  
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other  
US-09-778-927A-64

Query Match 71.1%; Score 32; DB 9; Length 279;  
Best Local Similarity 55.6%; Pred. No. 5.7e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTYDLFFV 9  
DB 233 YMSYHLFTIV 241

RESULT 40  
US-10-336-603A-48  
; Sequence 48, Application US/10336603A  
; Publication No. US20040072997A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-533A  
; CURRENT APPLICATION NUMBER: US/10/336, 603A  
; CURRENT FILING DATE: 2003-01-03  
; PRIOR APPLICATION NUMBER: 09/746,491  
; PRIOR FILING DATE: 2000-12-20  
; PRIOR APPLICATION NUMBER: 10/055,569  
; PRIOR FILING DATE: 2001-10-26  
; NUMBER OF SEQ ID NOS: 169  
; SOFTWARE: CuraSequlet version 0.1  
; SEQ ID NO 48  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-336-603A-48

Query Match 71.1%; Score 32; DB 15; Length 288;  
Best Local Similarity 55.6%; Pred. No. 5.9e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTYDLFFV 9  
DB 219 YMSYHLFTIV 227

Search completed: January 12, 2005, 20:48:28  
Job time : 95 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:46:14 ; Search time 16.4 Seconds  
(without alignments)  
52.802 Million cell updates/sec

Title: US-09-870-216C-9  
Perfect score: 45  
Sequence: 1 FLSYDLFV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	91.1	247	2	A82921 protein phosphatase
2	37	82.2	259	2	C89964 truncated transpos
3	36	80.0	191	2	F72277 hypothetical prote
4	34	75.6	95	2	A10315 conserved hypochet
5	34	75.6	412	2	C64712 ubiquinol-cytochro
6	34	75.6	412	2	D71803 ubiquinol-cytochro
7	34	75.6	452	2	C72295 hypothetical prote
8	33	73.3	70	2	S01213 NADH2 dehydrogenas
9	33	73.3	203	2	I40517 epig protein - Bac
10	33	73.3	274	2	T39166 inorganic phosphat
11	33	73.3	305	2	G86852 oxidoreductase ysj
12	33	73.3	333	2	T17712 hypothetical prote
13	33	73.3	370	2	T47504 mitochondrial prote
14	33	73.3	370	2	S40469 mitogen-activated
15	33	73.3	559	2	S62503 inorganic phosphat
16	33	73.3	572	2	T50404 probable inorganic
17	33	73.3	583	2	T41275 hypothetical prote
18	33	73.3	849	2	T20422 hypothetical prote
19	32	71.1	137	2	A83852 hypothetical prote
20	32	71.1	288	2	T78556 membrane glycoprot
21	32	71.1	309	2	G69796 lactose permease h
22	32	71.1	346	2	B90243 aminomethyltransfe
23	32	71.1	397	2	B98293 mannamate dehydrat
24	32	71.1	397	2	AG2990 mannamate dehydrat
25	32	71.1	413	2	D86920 probable UDP-galac
26	32	71.1	586	2	T19075 hypothetical prote
27	32	71.1	614	2	T05387 hypothetical prote
28	32	71.1	726	1	S73915 vitellence-associat
29	32	71.1	800	2	T38032 hypothetical prote

30	32	71.1	855	2	G86189 hypothetical prote
31	32	71.1	865	2	A85032 hypothetical prote
32	32	71.1	879	2	S49910 chloroplast outer
33	32	71.1	1503	2	T01098 chloroplast outer
34	32	71.1	1778	2	JT0382 apolipoprotein B -
35	32	71.1	2629	2	I46569 apolipoprotein B -
36	31	68.9	68	2	D86353 protein F2E2.10 (i
37	31	68.9	114	2	D98006 hypothetical prote
38	31	68.9	163	2	F81339 probable membrane
39	31	68.9	226	2	G95247 hypothetical prote
40	31	68.9	229	2	E90271 hypothetical prote
41	31	68.9	246	2	I51325 proteolipid protei
42	31	68.9	246	2	D98112 hypothetical prote
43	31	68.9	267	2	S39364 myfB protein precu
44	31	68.9	290	2	A11597 Lactococcus lactis
45	31	68.9	297	2	S66096 yabT protein - Bac

## ALIGNMENTS

```
RESULT 1
A82921 protein phosphatase U0215 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: A82921
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A:Reference number: A82870
A:Accession: A82921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <GUA>
A:Cross-references: GB:AF002120; GB:AF222894; NID:G6899167; PIDN:AAF30623.1; GSPDB:GN001
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: ptc1; U0215
A:Genetic code: SGC3

Query Match          91.1%; Score 41; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 FLSYDLFV 8
Db 176 FLSYDLFV 183

RESULT 2
C89964 truncated transposase [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C89964
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsum, K.
lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89964
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <KUR>
A:Cross-references: UNIPROT:Q09T82; GB:BA000018; PID:G13701580; PIDN:BAB42873.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: truncated-SA

Query Match          82.2%; Score 37; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.2;
```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLFV 7  
 |||||  
 Db 12 FLSYDLFV 18

RESULT 3  
 F72277  
 hypothetical protein - Thermotoga maritima (strain MSB8)

C/Species: Thermotoga maritima  
 C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C/Accession: F72277  
 C/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.  
 Nature 399, 323-329, 1999  
 A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
 A/Reference number: A72200; PMID:99287316; PMID:10360571  
 A/Accession: F72277  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-191 <ARN>  
 A/Cross-references: UNIPROT:Q9X0X9; GB:AE001780; GB:AE000512; NID:G4981796; PIDN:AAD3632  
 A/Experimental source: strain MSB8  
 C/Genetics:  
 A/Gene: TM1252

Query Match 80.0%; Score 36; DB 2; Length 191;  
 Best Local Similarity 66.7%; Pred. No. 8.5;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 |||||  
 Db 44 FLSYDLFV 52

## RESULT 4

conserved hypothetical protein YPO2589 [imported] - Yersinia pestis (strain CO92)

A/Accession: F72277  
 C/Species: Yersinia pestis  
 C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: A10315  
 R.Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; H., M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, Nature 413, 523-527, 2001  
 A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A/Reference number: AB0001; PMID:11586360  
 A/Accession: A10315  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-95 <KOR>  
 A/Cross-references: UNIPROT:Q8ZDH9; GB:AL590842; PIDN:CA91389.1; PID:G15980575; GSPDB:C  
 C/Genetics:  
 A/Gene: YPO2589

Query Match 75.6%; Score 34; DB 2; Length 95;  
 Best Local Similarity 85.7%; Pred. No. 11;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
 |||||  
 Db 36 LSYDLFV 42

## RESULT 5

C64712  
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Helicobacter pylori (strain  
 C/Species: Helicobacter pylori  
 C/Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
 C/Accession: C64712  
 R.Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khajak, H.G.; Glodet, A.; McKenna, son, J.D.; Kelley, J.M.; Cotton, M.D.; Weldman, J.M.; Fujii, C.; Bowman, C.; Watney, L. Nature 388, 539-547, 1997  
 A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.; A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
 A/Reference number: A64520; PMID:9794467; PMID:9252185  
 A/Accession: C64712  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-412 <TOM>

A/Cross-references: UNIPROT:Q26064; GB:AE000652; GB:AE000511; NID:G2314720; PIDN:AAD0857  
 C/Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyaninol  
 C/Keywords: electron transfer; heme; iron; metalloprotein; oxidoreductase  
 F/21-370/Domain: cytochrome b homology <CBH>  
 F/21-223/Domain: cytochrome b6 homology <CB6>  
 F/250-370/Domain: plastocyaninol-plastocyanin reductase 17k protein homology <17k>  
 F/94,195/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted  
 F/108,210/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 |||||  
 Db 253 FLSYDLFV 261

## RESULT 6

D71803  
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Helicobacter pylori (strain

C/Species: Helicobacter pylori  
 A/Variety: strain J99  
 C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
 C/Accession: D71803  
 R.Aim, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doly, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
 A/Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
 A/Reference number: A71800; PMID:9910557; PMID:9923682  
 A/Accession: D71803  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-412 <ARN>  
 A/Cross-references: UNIPROT:Q9ZJ55; GB:AE001568; GB:AE001439; NID:G4156083; PIDN:AAD0704  
 A/Experimental source: strain J99  
 C/Genetics:

A/Gene: petB  
 C/Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyaninol  
 C/Keywords: electron transfer; heme; iron; metalloprotein; oxidoreductase  
 F/21-370/Domain: cytochrome b homology <CYB>  
 F/21-223/Domain: cytochrome b6 homology <CB6>  
 F/250-370/Domain: plastocyaninol-plastocyanin reductase 17k protein homology <17k>  
 F/94,195/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted  
 F/108,210/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 75.6%; Score 34; DB 2; Length 412;  
 Best Local Similarity 77.8%; Pred. No. 45;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 |||||  
 Db 253 FLSYDLFV 261

## RESULT 7

C72295  
 hypothetical protein TM105 - Thermotoga maritima (strain MSB8)

C/Species: Thermotoga maritima  
 C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C/Accession: C72295  
 C/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.

C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: C72295  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-452 <ARN>  
A:Cross-references: UNIPROT:Q9X019; GB:AB001769; GB:AB000512; NID:g4981639; PIDN:AA03618  
C:Genetic8:  
C:Superfamily: [Nife]-hydrogenase-3-type complex, large membrane subunit  
Query Match 75.6%; Score 34; DB 2; Length 452;  
Best Local Similarity 85.7%; Pred. No. 49;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 FLSYDLFV 7  
DB 102 FVSYDLF 108  
RESULT 8  
501213  
NADH dehydrogenase (ubiquinone) (BC 1.6.5.3) chain 3 - brine shrimp mitochondrion (frag  
C:Species: mitochondrion Artemia sp. (brine shrimp)  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 03-Jun-2002  
C:Accession: S01213  
R:Batuecas, B.; Garrese, R.; Calleja, M.; Valverde, J.R.; Marco, R.  
Nucleic Acids Res. 16, 6515-6529, 1988  
A:Title: Genome organization of Artemia mitochondrial DNA.  
A:Reference number: S01207; MUID:88289417; PMID:3135541  
A:Accession: S01213  
A:Molecule type: DNA  
A:Residues: 1-31;32-70 <BAT>  
A:Cross-references: EMBL:X07665  
A:Note: the translation of residues 1-8 is inconsistent with the nucleotide sequence  
C:Genetics:  
A:Genome: mitochondrion  
A:Genetic code: SGC4  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 3  
C:Keywords: membrane-associated complex; mitochondrion; NMD; oxidative phosphorylation;  
Query Match 73.3%; Score 33; DB 2; Length 70;  
Best Local Similarity 75.0%; Pred. No. 13;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLSYDLFV 8  
DB 33 FLSYDLFV 40  
RESULT 9  
140517  
spac protein - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 09-Jul-2004  
C:Accession: 140517  
R:Klein, C.; Entian, K.D.  
Appl. Environ. Microbiol. 60, 2793-2801, 1994  
A:Title: Genes involved in self-protection against the antibiotic subtilin produced by  
A:Reference number: 140511; MUID:94368094; PMID:8088823  
A:Accession: 140517  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-203 <RES>  
A:Cross-references: UNIPROT:Q45405; EMBL:U09819; NID:g2702240; PIDN:AA891595.1; PID:g595  
C:Genetics:  
A:Gene: spac  
Query Match 73.3%; Score 33; DB 2; Length 203;  
Best Local Similarity 62.5%; Pred. No. 36;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 2 LSYDLFV 9  
DB 78 LSYDLFV 85  
RESULT 10  
739166  
Inorganic phosphate transporter [similarity] - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jun-2000  
C:Accession: T50381; T39166  
R:Connor, R.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, February 1998  
A:Reference number: Z21832  
A:Accession: T50381  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-274 <CO2>  
A:Cross-references: EMBL:AL021815; PIDN:CA16994.1; GSPDB:GN00067; SPDB:SPBC8E4.01c  
A:Experimental source: strain 972h-; cosmid c8E4  
C:Genetics:  
A:Gene: SPDB:SPBC8E4.01c  
A:Map position: 2  
C:Superfamily: probable inorganic phosphate transport protein PHO84  
Query Match 73.3%; Score 33; DB 2; Length 274;  
Best Local Similarity 71.4%; Pred. No. 48;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 3 SYDLFV 9  
DB 59 SYDLFV 65  
RESULT 11  
G86852  
oxidoreductase ysjB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: G86852  
R:Bojorin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
A:Reference number: A86825; MUID:21235186; PMID:11337471  
A:Accession: G86852  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-305 <STO>  
A:Cross-references: UNIPROT:Q9CEL4; GB:AB005176; PID:g12724851; PIDN:AAK05921.1; GSPDB:G  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: ysjB  
Query Match 73.3%; Score 33; DB 2; Length 305;  
Best Local Similarity 62.5%; Pred. No. 53;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FLSYDLFV 8  
DB 155 FLMDYDLF 162  
RESULT 12  
T17712  
Hypothetical protein A222R - Chlorella virus PBCV-1  
C:Species: Chlorella virus PBCV-1  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T17712  
R:Graves, M.V.; Van Etten, J.L.  
submitted to the EMBL Data Library, May 1999  
A:Reference number: Z18806

```

A/Accession: T17712
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-333 <GRA>
A/Cross-references: UNIPROT:Q84542, EMBL:U42580, NID:g4028986, PIDN:AAC96590.1
A/Experimental source: specific host Chlorella strain NC64A
C/Genetics:
A/Note: A222R
C/Superfamily: Chlorella virus PBCV-1 hypothetical protein A222R

Query Match          73.3%; Score 33; DB 2; Length 333;
Best Local Similarity 75.0%; Pred. No. 58;
Matches 6; Conservative 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 8
Db 254 FFSFDLFV 261

RESULT 13
T47504
mitogen-activated protein kinase 3 - Arabidopsis thaliana
N/Alternate names: protein P9K21.220
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C/Accession: T47504
R/Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke, K.
submitted to the Protein Sequence Database, February 2000
A/Reference number: Z24467
A/Accession: T47504
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-370 <JOR>
A/Cross-references: UNIPROT:Q39023; EMBL:AL138657
A/Experimental source: cultivar Columbia; BAC clone P9K21
C/Genetics:
A/Map position: 3
A/Introns: 53/2; 96/3; 142/3; 253/3; 315/1
A/Note: P9K21.220
C/Superfamily: kinase-related transforming protein; protein kinase homology

Query Match          73.3%; Score 33; DB 2; Length 370;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLF 7
Db 21 FISYDLF 27

RESULT 14
S40469
mitogen-activated protein kinase 3 (EC 2.7.1.-) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 06-Oct-1994 #sequence_revision 19-Jan-1996 #text_change 09-Jul-2004
C/Accession: S40469
R/Mizoguchi, T.; Hayashida, N.; Yamaguchi-Shinozaki, K.; Kamada, H.; Shinozaki, K.
FEBS Lett. 336, 440-444, 1993
A/Title: ATPKs: a gene family of plant MAP kinases in Arabidopsis thaliana.
A/Reference number: S40469; MUID:94109583; PMID:8282107
A/Accession: S40469
A/Molecule type: mRNA
A/Residues: 1-370 <MTZ>
A/Cross-references: UNIPROT:Q39023; EMBL:D21839; NID:g457397; PIDN:BA04866.1; PID:g4573
C/Genetics:
A/Genes: MPK3
C/Superfamily: kinase-related transforming protein; protein kinase homology
C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F15-324/Domain: protein kinase homology <KIN>
F144-52/Region: protein kinase ATP-binding motif

Query Match          73.3%; Score 33; DB 2; Length 370;
Best Local Similarity 71.4%; Pred. No. 64;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLF 7
Db 21 FISYDLF 27

RESULT 15
S62503
inorganic phosphate transporter - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C/Accession: T38287; S62503
R/Mblett, D.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, October 1995
A/Reference number: Z21783
A/Accession: T38287
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-559 <NI2>
A/Cross-references: UNIPROT:Q09852; EMBL:Z64354; NID:g1039338; PIDN:CA01247.1; PID:g103
C/Genetics:
A/Genes: SPDB:SPAC23D3.12
A/Map position: 1K
C/Superfamily: probable inorganic phosphate transport protein PHO84

Query Match          73.3%; Score 33; DB 2; Length 559;
Best Local Similarity 71.4%; Pred. No. 96;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 SYDLFV 9
Db 57 SYDLFI 63

RESULT 16
T50404
probable inorganic phosphate transporter [imported] - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C/Accession: T50404
R/Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 2000
A/Reference number: Z25068
A/Accession: T50404
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-572 <RIE>
A/Cross-references: UNIPROT:Q42885; EMBL:AL137099; PIDN:CA06656.1; GSPDB:GN00067; SPDB:
A/Experimental source: strain 972h(-); clone pl p4G3
C/Genetics:
A/Genes: SPBC854.01c; SPDB:SPBP4G3.01
A/Map position: 2
C/Superfamily: probable inorganic phosphate transport protein PHO84

Query Match          73.3%; Score 33; DB 2; Length 572;
Best Local Similarity 71.4%; Pred. No. 98;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 SYDLFV 9
Db 59 SYDLFI 65

RESULT 17
T41275
probable inorganic phosphate transporter - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T41275
R/Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckert, G.
submitted to the EMBL Data Library, March 1999

```

A/Reference number: 221962  
 A/Accession: T41275  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-583 <LVN>  
 A/Cross-references: UNIPROT:Q9Y7Q9; EMBL:AL049567; PDB:1CAB40203.1; GSPDB:GN00068; SPDB:  
 A/Experimental source: strain 972h-; cosmid C2H8  
 C/Genetics:  
 A/Map position: 3  
 A/Suprafamily: probable inorganic phosphate transport protein PHO84

Query Match 73.3%; Score 33; DB 2; Length 583;  
 Best Local Similarity 71.4%; Pred. No. 1e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYDLFV 9  
 Db 53 SYDLFV 59

RESULT 18  
 T20422  
 hypochelical protein E02H4.4 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T20422  
 R/Barlow, K. submitted to the EMBL Data Library, November 1995  
 A/Reference number: Z19273  
 A/Accession: T20422  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-849 <WLL>  
 A/Cross-references: UNIPROT:Q19040; EMBL:Z68003; PDB:1CAA91977.1; GSPDB:GN00028; CESP:ED  
 A/Experimental source: clone E02H4  
 C/Genetics:  
 A/Map position: X  
 A/Intons: 19/1; 69/3; 100/2; 135/1; 234/1; 337/3; 380/3; 406/2; 441/2; 509/1; 576/3; 58

Query Match 73.3%; Score 33; DB 2; Length 849;  
 Best Local Similarity 55.6%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9  
 Db 315 FTHDIFV 323

RESULT 19  
 A83852  
 hypochelical protein BH1617 [imported] - Bacillus halodurans (strain C-125)  
 C/Species: Bacillus halodurans  
 C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C/Accession: A83852  
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A/Reference number: A83650; MUID:2051582; PMID:11058132  
 A/Accession: A83852  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-137 <STO>  
 A/Cross-references: UNIPROT:Q9KCF4; GB:AP001512; GB:BA000004; NID:910174030; PDB:1BAB053  
 A/Experimental source: strain C-125  
 C/Genetics:  
 A/Suprafamily: inner membrane protein C-125

Query Match 71.1%; Score 32; DB 2; Length 137;  
 Best Local Similarity 75.0%; Pred. No. 38;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 8  
 Db 59 FLSYDLFV 66

RESULT 20  
 I78556  
 membrane glycoprotein M6 - mouse  
 C/Species: Mus sp. (mouse)  
 C/Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 05-Nov-1999  
 C/Accession: I78556  
 R/Ryan, Y.; Lagenaar, C.; Narayanan, V.  
 Neuron 11, 423-431, 1993  
 A/Title: Molecular cloning of M6: identification of a PLP/DW20 gene family.  
 A/Reference number: I58150; MUID:9400809; PMID:8398137  
 A/Accession: I78556  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-288 <RES>  
 A/Cross-references: GB:S65737; NID:9425586; PDB:1AAB28351.1; PDB:1G425587  
 C/Suprafamily: myelin proteolipid protein  
 C/Keywords: glycoprotein

Query Match 71.1%; Score 32; DB 2; Length 288;  
 Best Local Similarity 55.6%; Pred. No. 79;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9  
 Db 219 FLSYDLFV 227

RESULT 21  
 G69796  
 lactose permease homolog yepF - Bacillus subtilis  
 C/Species: Bacillus subtilis  
 C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C/Accession: G69796  
 R/Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berti  
 C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carer, N.M.; Cho  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabbet, C.; Ferrari, B.  
 Nature 390, 249-256, 1997  
 A/Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler  
 lech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullio, M.F.  
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
 A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
 Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon  
 A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror  
 akenuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A/Reference number: A69580; MUID:98044033; PMID:9384377  
 A/Accession: G69796  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-309 <KUN>  
 A/Cross-references: UNIPROT:O31519; GB:Z99107; GB:AL009126; NID:92632866; PDB:1CAB12517.  
 A/Experimental source: strain 168  
 C/Genetics:  
 A/Suprafamily: inner membrane protein upgA

Query Match 71.1%; Score 32; DB 2; Length 309;  
 Best Local Similarity 85.7%; Pred. No. 85;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 7  
 Db 49 FLSYDLFV 55

## RESULT 22

B90243  
aminomethyltransferase [imported] - *Sulfolobus solfataricus*  
C:Species: *Sulfolobus solfataricus*  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C/Accession: B90243  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, H.  
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: *Sulfolobus solfataricus* complete genome.  
A:Reference number: A99139  
A:Accession: B90243  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-346 <KUR>  
A:Cross-references: UNIPROT:Q97Z17; GB:AE006641; NID:G13814101; PIDN:AAK41201.1; GSPDB:G  
C:Genetics:  
A:Gene: SS00919  
C:Superfamily: aminomethyltransferase

Query Match 71.1%; Score 32; DB 2; Length 346;  
Best Local Similarity 44.4%; Pred. No. 95;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLSYDLFV 9  
|:|:|:|:  
Db 167 FVGVDVFLI 175

## RESULT 23

B98293  
mannonate dehydratase (d-mannonate hydrolase) [imported] - *Agrobacterium tumefaciens* (str  
C:Species: *Agrobacterium tumefaciens*  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C/Accession: B98293  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2333-2328, 2001  
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: B98293  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-397 <KUR>  
A:Cross-references: UNIPROT:Q8UA46; GB:AE007870; PIDN:AAK89868.1; PID:G15159811; GSPDB:G  
C:Genetics:  
A:Gene: AGR\_L\_2605  
A:Map position: linear chromosome

Query Match 71.1%; Score 32; DB 2; Length 397;  
Best Local Similarity 44.4%; Pred. No. 1.1e+02;  
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLSYDLFV 9  
|:|:|:|:  
Db 133 FVAYDVFIL 141

## RESULT 24

AG2990  
mannonate dehydratase [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)  
C:Species: *Agrobacterium tumefaciens*  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C/Accession: AG2990  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McChie  
i Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreppan, W.; Perry, M.; Gordon-Kamm,  
ster, B.W.  
A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AG2990  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-397 <KUR>  
A:Cross-references: UNIPROT:Q8UA46; GB:AE008689; PIDN:AAK44341.1; PID:G17741933; GSPDB:G  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: uxvA  
A:Map position: linear chromosome

Query Match 71.1%; Score 32; DB 2; Length 397;  
Best Local Similarity 44.4%; Pred. No. 1.1e+02;  
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 1 FLSYDLFV 9  
|:|:|:|:  
Db 133 FVAYDVFIL 141

## RESULT 25

D86920  
probable UDP-galactopyranose mutase [imported] - *Mycobacterium leprae*  
C:Species: *Mycobacterium leprae*  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C/Accession: D86920  
R:Coile, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq  
A>Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86909; MUID:21128732; PMID:11234002  
A:Accession: D86920  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-413 <STO>  
A:Cross-references: UNIPROT:Q9CDB8; GB:AL450380; NID:G13092474; PIDN:CAK29600.1; GSPDB:G  
C:Genetics:  
A:Gene: glf  
C:Superfamily: Methanobacterium thermoautotrophicum UDP-galactopyranose mutase

Query Match 71.1%; Score 32; DB 2; Length 413;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 LSYDLFV 9  
|:|:|:|:  
Db 3 VSPDLFV 10

## RESULT 26

T19075  
hypothetical protein C08B6.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T19075  
R:Milkinson, J.  
submitted to the EMBL Data Library, May 1996  
A:Reference number: Z19070  
A:Accession: T19075  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-586 <WIL>  
A:Cross-references: UNIPROT:Q17813; EMBL:Z72502; PIDN:CAA96584.1; GSPDB:GN00023; CESP:CO  
A:Experimental source: clone C08B6  
C:Genetics:  
A:Gene: CESP:C08B6.1  
A:Map position: 5  
A:Introns: 52/3; 128/1; 245/1; 419/1; 447/3  
C:Superfamily: glucuronosyltransferase

Query Match 71.1%; Score 32; DB 2; Length 586;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;



Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 FLSYDLFV 9  
| | | | |  
Db 544 FYSYDLIV 552

## RESULT 27

hypothetical protein F16G20.190 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T05387  
R:Bevan, M.; Watson, M.D.; Gibbons, T.; Bartley, J.; Bancroft, I.; Mewes, H.W.; Mayer, K.  
submitted to the Protein Sequence Database, August 1998  
A:Reference number: 215413  
A:Accession: T05387  
A:Molecule type: DNA  
A:Residues: 1-614 <BEV>  
A:Cross-references: UNIPROT:O81745; EMBL:AL031326  
A:Experimental source: cultivar Columbia; BAC clone F16G20  
C:Genetics:  
A:Map position: 4  
A:Introns: 14/1; 61/3; 121/1; 163/3; 324/3; 402/3  
A:Note: F16G20.190  
C:Superfamily: Arabidopsis hypothetical protein F13M22.23

Query March 71.1%; Score 32; DB 2; Length 614;  
Best Local Similarity 62.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 9  
| | | | |  
Db 591 LAYDLVY 598

## RESULT 28

virulence-associated protein vacB homolog - Mycoplasma pneumoniae (strain ATCC 29342)  
N:Alternate names: hypothetical protein K04\_crf726  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: S73915  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A:Reference number: S73327; MUID:97105885; PMID:8948633  
A:Accession: S73915  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-726 <HIM>  
A:Cross-references: UNIPROT:P75529; EMBL:AE000057; GB:U00089; NID:g1674279; PIDN:AAB9623  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Gene: vacB  
A:Genetic code: SGC3  
C:Superfamily: virulence-associated protein vacB homolog

Query March 71.1%; Score 32; DB 1; Length 726;  
Best Local Similarity 71.4%; Pred. No. 2e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYDLFV 9  
| | | | |  
Db 537 SYDLFV 543

## RESULT 29

T38032  
hypothetical protein SPAC1B3.13 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T38032  
R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, August 1997  
A:Reference number: 221763  
A:Accession: T38032

A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-800 <MUR>  
A:Cross-references: UNIPROT:O13878; EMBL:Z98598; PIDN:CAB11248.1; GSPDB:GN00066; SPDB:SP  
A:Experimental source: strain 972h; cosmid c1B3  
C:Genetics:  
A:Gene: SPDB:SPAC1B3.13  
A:Map position: 1  
A:Introns: 39/2

Query March 71.1%; Score 32; DB 2; Length 800;  
Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 8  
| | | | |  
Db 42 FLTDTFV 49

## RESULT 30

hypothetical protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: G86189  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chao, M.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hultzer, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzilli, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: G86189  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-855 <STO>  
A:Cross-references: UNIPROT:Q9SYU9; GB:AB005172; NID:g4836908; PIDN:AAD30610.1; GSPDB:GN  
C:Genetics:  
A:Map position: 1

Query March 71.1%; Score 32; DB 2; Length 855;  
Best Local Similarity 85.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SYDLFV 9  
| | | | |  
Db 734 SYDLVY 740

## RESULT 31

A85032  
hypothetical protein ATAG02510 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: A85032  
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: A85032  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-865 <STO>  
A:Cross-references: UNIPROT:O22774; GB:NC\_001268; NID:g7269011; PIDN:CAB80744.1; GSPDB:GN

C:Genetics:  
A:Gene: ATG02510  
A:Map position: 4

Query Match 71.1%; Score 32; DB 2; Length 865;  
Best Local Similarity 85.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
Db 358 LSYDLFV 364

## RESULT 32

S49910  
Chloroplast outer envelope protein OEP86 precursor - garden pea  
N:Alternate names: chloroplast import-associated protein IAP86, GTP-binding  
C:Species: Pisum sativum (garden pea)  
C:Date: 26-Jan-1995 #sequence revision 10-Feb-1995 #ext\_change 09-Jul-2004  
C:Accession: S49910; A55386; A55171

R:Hirsch, S.; Soll, J.  
submitted to the EMBL Data Library, March 1994  
A:Reference number: S49910

A:Accession: S49910

A:Molecule type: mRNA

A:Residues: 1-879 <HIR>

A:Cross-references: UNIPROT:Q41010; EMBL:Z31581; NID:g599957; PID:g599958

R:Hirsch, S.; Muckel, E.; Heemeyer, F.; von Heijne, G.; Soll, J.

Science 266, 1989-1992, 1994

A:Title: A receptor component of the chloroplast protein translocation machinery.

A:Reference number: A55386; NID:95099324; PMID:7801125

A:Accession: A55386

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-137, 'Q', 139-875, 877-879 <HI2>

R:Kessler, F.; Blobel, G.; Patel, H.A.; Schnell, D.J.

Science 266, 1035-1039, 1994

A:Title: Identification of two GTP-binding proteins in the chloroplast protein import ma

A:Reference number: A55171; NID:95063938; PMID:7973656

A:Accession: A55171

A:Molecule type: mRNA

A:Residues: 1-372, 'EQ', 376-879 <KES>

A:Cross-references: GB:I36857; NID:9576508; PID:AAA3276.1; PID:9576509

C:Keywords: chloroplast; GTP binding; membrane protein; nucleotide binding; P-loop

F:243-252/Region: nucleotide-binding motif A (P-loop)

Query Match 71.1%; Score 32; DB 2; Length 879;  
Best Local Similarity 85.7%; Pred. No. 2.4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
Db 379 LSYDLFV 385

## RESULT 33

T01098  
Chloroplast outer envelope protein OEP86 homolog T10P11.19 - Arabidopsis thaliana  
N:Alternate names: chloroplast import-associated protein, GTP-binding; protein T14P8.24  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence revision 12-Feb-1999 #ext\_change 09-Jul-2004

C:Accession: T01098; T01299

R:Kaplan, N.; Johnson, D.; Schütz, K.; Gnoj, L.; Hoffman, U.; Tüll, S.; de la Bastide, M.

h.; Martienssen, R.; Chen, E.Y.; Wilson, R.; McCombie, W.R.

submitted to the EMBL Data Library, November 1998

A:Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.

A:Reference number: Z14248

A:Accession: T01098

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1503 <KAP>

A:Cross-references: UNIPROT:O81283; EMBL:AC002330; NID:g2262135; PID:g3892053

A:Experimental source: cultivar Columbia

R:Kalicki, J.; Elliott, G.; Cloud, J.  
submitted to the EMBL Data Library, May 1998  
A:Description: The sequence of A. thaliana T14P8.

A:Reference number: Z14290

A:Accession: T01299

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1503 <KAL>

A:Cross-references: EMBL:AF069298; NID:g3193282; PID:g3193301

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:introns: 22/1

A>Note: T10P11.19; T14P8.24

C:Keywords: chloroplast; GTP binding; membrane protein; nucleotide binding; P-loop

F:862-869/Region: nucleotide-binding motif A (P-loop)

Query Match 71.1%; Score 32; DB 2; Length 1503;  
Best Local Similarity 85.7%; Pred. No. 4e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
Db 996 LSYDLFV 1002

## RESULT 34

JT0382  
apolipoprotein B - pig (fragments)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #ext\_change 09-Jul-2004

C:Accession: JT0382; I46567; I46568

R:Meda, N.; Ebert, D.L.; Doers, T.M.; Newman, M.; Hasler-Rapacz, J.; Attie, A.D.; Rapacz

Gene 70, 213-229, 1988

A:Title: Molecular genetics of the apolipoprotein B gene in pigs in relation to atherosc

A:Reference number: JT0382; NID:89108006; PMID:2905687

A:Accession: JT0382

A:Molecule type: DNA

A:Residues: 1-1778 <MAB>

A:Cross-references: UNIPROT:Q29020; UNIPROT:Q29433

R:Meda, N.; Ebert, D.L.; Doers, T.M.; Newman, M.; Hasler-Rapacz, J.O.; Attie, A.D.; Rap

Gene 69, 213-229, 1988

A:Title: Molecular genetics of the apolipoprotein B gene in pigs in relation to atherosc

A:Reference number: I46567

A:Accession: I46567

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-8, 'S', 10-238 <MA2>

A:Cross-references: GB:M26447; NID:g164367; PIDN:AAA30996.1; PID:g164369

A:Accession: I46568

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 239-1778 <MA3>

A:Cross-references: GB:M26447; NID:g164367; PIDN:AAA30997.1; PID:g164370

C:Comment: Apolipoprotein B is the predominant protein component of the low-density lipc

C:Genetics:

A:Gene: apoB

A:introns: 39/3; 88/3; 159/2; 238/3; 1133/1; 1171/2; 1232/3

A>Note: the list of introns may be incomplete

C:Superfamily: apolipoprotein B

C:Keywords: atherosclerosis; cholesterol metabolism; LDL; lipid binding; lipoprotein

Query Match 71.1%; Score 32; DB 2; Length 1778;  
Best Local Similarity 77.8%; Pred. No. 4.7e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
Db 1126 FLSYDLFV 1134

## RESULT 35

I46569

apolipoprotein B - pig (fragment)  
 C/Species: Sus scrofa domestica (domestic pig)  
 C/Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 09-Jul-2004  
 C/Accession: I46569  
 R/Putrell, C.; Maeda, N.; Eberly, D.L.; Kaiser, M.; Lund-Katz, S.; Scurley, S.L.; Kodoyia  
 J. Lipid Res. 34, 1333-1335, 1993  
 A/Title: Nucleotide sequence encoding the carboxyl-terminal half of apolipoprotein B  
 A/Reference number: I46569; MUID:94014802; PMID:8409766  
 A/Accession: I46569  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-2629 <PUR>  
 A/Cross-references: UNIPROT:Q29021; GB:L11235; NID:9164371; PIDN:AAA74655.1; PID:9951375  
 C/Genetics:  
 A/Genes: APOB  
 A/Supernotes: 1984/1; 2022/2; 2083/3  
 C/Superfamily: apolipoprotein B

Query Match 71.1%; Score 32; DB 2; Length 2629;  
 Best Local Similarity 77.8%; Pred. No. 6.9e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 FLSYDLFV 9  
 Db 1977 FLSYDLNV 1985

RESULT 36  
 D86353  
 protein F282.10 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C/Accession: D86353  
 R/Theologian, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.R.; Hughes, B.; Hutzar, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A/Reference number: A86141; MUID:21016719; PMID:11130712  
 A/Accession: D86353  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-68 <STO>  
 A/Cross-references: UNIPROT:Q9JLM56; GB:AE005172; NID:99280678; PIDN:AAE86547.1; GSPDB:GN  
 C/Genetics:  
 A/Genes: F282.10  
 A/Map position: 1

Query Match 68.9%; Score 31; DB 2; Length 68;  
 Best Local Similarity 75.0%; Pred. No. 31;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 LSYDLFV 9  
 Db 24 LRVDVFFV 31

RESULT 37  
 D98006  
 hypothetical protein lack-truncation [imported] - Streptococcus pneumoniae (strain R6)  
 C/Species: Streptococcus pneumoniae  
 C/Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
 C/Accession: D98006  
 R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E  
 e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
 y, P.; Sun, P.M.; Winkler, M.E.  
 J. Bacteriol. 183, 5709-5717, 2001  
 A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
 A/Reference number: A97872; MUID:21429245; PMID:11544234  
 A/Accession: D98006  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-114 <KUR>  
 A/Cross-references: UNIPROT:Q8DPN8; GB:AE007317; PIDN:AAK99880.1; PID:915458700; GSPDB:G  
 C/Genetics:  
 A/Genes: lack-truncation

Query Match 68.9%; Score 31; DB 2; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LSYDLF 7  
 Db 19 LSYDLF 24

RESULT 38  
 F81339  
 probable membrane protein Cj0691 [imported] - Campylobacter jejuni (strain NCTC 11168)  
 C/Species: Campylobacter jejuni  
 C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
 C/Accession: F81339  
 R/Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling  
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVlier, A.; Whitehead, S.; Barrel  
 Nature 403, 665-668, 2000  
 A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
 A/Reference number: A81250; MUID:20150912; PMID:10688204  
 A/Accession: F81339  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-163 <PAR>  
 A/Cross-references: UNIPROT:Q9PPL6; GB:AL139076; GB:AL111168; NID:96968128; PIDN:CA87296  
 A/Experimental source: serotype O2, strain NCTC 11168  
 C/Genetics:  
 A/Genes: Cj0691  
 C/Superfamily: Campylobacter jejuni probable membrane protein Cj0691

Query Match 68.9%; Score 31; DB 2; Length 163;  
 Best Local Similarity 55.6%; Pred. No. 72;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 FLSYDLFV 9  
 Db 95 FLSYDLFV 103

RESULT 39  
 G95247  
 hypothetical protein SP2117 [imported] - Streptococcus pneumoniae (strain TIGR4)  
 C/Species: Streptococcus pneumoniae  
 C/Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 09-Jul-2004  
 C/Accession: G95247  
 R/Retelien, H.; Nelson, K.R.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzaple,  
 neon, T.; Hickey, E.K.; Holt, I.E.  
 Science 293, 498-506, 2001  
 A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
 A/Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
 A/Reference number: A95000; MUID:21357209; PMID:11463916  
 A/Accession: G95247  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-226 <KUR>  
 A/Cross-references: UNIPROT:Q9TND3; GB:AE005672; PIDN:AAK76176.1; PID:914973629; GSPDB:G  
 A/Experimental source: strain TIGR4  
 C/Genetics:  
 A/Genes: SP2117

Query Match 68.9%; Score 31; DB 2; Length 226;  
 Best Local Similarity 77.8%; Pred. No. 99;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9

Db 113 FLSYDLFV 121

RESULT 40

E90271 hypothetical protein SS01180 [imported] - Sulfolobus solfataricus transposon IS01439

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C/Accession: E90271

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Chan-Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R. arrett, R.A.; Kagan, M.A.; Sersen, C.W.; Van der Oost, J.

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A9139

A/Accession: E90271

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-229 <KUR>

A/Cross-references: UNIPROT:Q97YX2; GB:AE006641; NID:913814372; PIDN:AAK41428.1; GSPDB:G

C/Genetics:

A/Gene: SS01180

Query Match 68.9%; Score 31; DB 2; Length 229;

Best Local Similarity 85.7%; Pred. No. 1e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8

Db 220 LSYDLFV 226

Search completed: January 12, 2005, 20:15:48  
Job time : 18.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:38:08 ; Search time 85 Seconds  
(without alignments)  
60.922 Million cell updates/sec

Title: US-09-870-216C-9  
Perfect score: 45  
Sequence: 1 FLSTYDLFV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_02.\*  
1: uniProt\_sprot.\*  
2: uniProt\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	91.1	247	2 Q9POS7	Q9POS7 ureaplasma
2	38	84.4	69	2 Q7BHL0	Q7BHL0 fusobacteri
3	38	84.4	69	2 AAS45303	AAS45303 fusobacte
4	37	82.2	259	2 Q99T82	Q99T82 streptococ
5	36	80.0	191	2 Q9X0X9	Q9X0X9 thermotoga
6	35	77.8	101	2 Q9D9A8	Q9D9A8 mus musculu
7	35	77.8	162	2 Q9NPI0	Q9NPI0 homo sapien
8	35	77.8	162	2 Q9D6G5	Q9D6G5 m mus muscu
9	35	77.8	162	2 AAS58237	AAS58237 mus muscu
10	35	77.8	192	2 Q812F8	Q812F8 plasmodium
11	35	77.8	203	2 Q84AF2	Q84AF2 bacillus st
12	35	75.6	35	2 Q882I6	Q882I6 pseudomonas
13	34	75.6	95	2 Q8ZDH9	Q8ZDH9 yersinia pe
14	34	75.6	95	2 AAS61372	AAS61372 yersinia
15	34	75.6	184	2 Q6L299	Q6L299 picophilus
16	34	75.6	382	2 Q9ZPF4	Q9ZPF4 arabidopsis
17	34	75.6	389	1 UKUA_BACTN	UKUA_BACTN bacteroides
18	34	75.6	394	1 UKUA_RHIL0	UKUA_RHIL0 rhizobium
19	34	75.6	409	1 Q7VHG1	Q7VHG1 helicobacte
20	34	75.6	411	2 Q7M7Q1	Q7M7Q1 wolfinella s
21	34	75.6	412	2 Q26064	Q26064 helicobacte
22	34	75.6	412	2 Q9ZJ55	Q9ZJ55 helicobacte
23	34	75.6	452	2 Q9X0I9	Q9X0I9 thermotoga
24	34	75.6	516	2 Q9ZWT3	Q9ZWT3 arabidopsis
25	34	75.6	590	2 Q8Y105	Q8Y105 ralsstonia s
26	34	75.6	618	2 Q7N1U9	Q7N1U9 gloeobacter
27	34	75.6	674	1 K8TE_HUMAN	K8TE_HUMAN homo sapie
28	34	75.6	729	2 Q91L83	Q91L83 white spot
29	34	75.6	924	2 Q6CFR0	Q6CFR0 yarrowia li
30	34	75.6	1559	2 Q8IDC3	Q8IDC3 plasmodium
31	33	73.3	70	1 NU3M_ARISA	NU3M_ARISA attemia sal

32	33	73.3	203	2 Q45405	Q45405 bacillus su
33	33	73.3	305	2 Q9CEL4	Q9CEL4 lactococcus
34	33	73.3	332	2 Q7ZHT6	Q7ZHT6 thermus the
35	33	73.3	332	2 AAS81742	AAS81742 thermus t
36	33	73.3	355	2 Q8RD34	Q8RD34 thermoaer
37	33	73.3	370	1 MPK3_ARATH	MPK3_ARATH arabidopsi
38	33	73.3	370	2 ANI15326	ANI15326 arabidops
39	33	73.3	432	2 Q84542	Q84542 parametium
40	33	73.3	479	2 Q9DHK3	Q9DHK3 yaba-like d
41	33	73.3	559	1 YAEI_SCHPO	YAEI_SCHPO schizosacch
42	33	73.3	572	1 YHDI_SCHPO	YHDI_SCHPO schizosacch
43	33	73.3	573	1 Q9Y7Q9	Q9Y7Q9 schizosacch
44	33	73.3	583	2 Q9Y7Q9	Q9Y7Q9 schizosacch
45	33	73.3	607	2 Q8A407	Q8A407 bacteroides

## ALIGNMENTS

RESULT 1					
ID	Q9POS7	PRELIMINARY;	PRT;	247 AA.	
AC	Q9POS7;				
DT	01-OCT-2000 (TREMBLrel. 15, Created)				
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)				
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)				
DE	Protein phosphatase.				
GN	Name=ptci; OrderedlocusNames=UU215;				
OS	Ureaplasma parvum (Ureaplasma urealyticum biotype 1).				
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.				
OX	NCBI_TaxID=134821;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 700970;				
RX	MEDLINE=20500219; PubMed=11048724;				
RA	Glass J.T., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,				
RA	Caesell G.H.;				
RT	"The complete sequence of the mucosal pathogen Ureaplasma				
RT	urealyticum.";				
RL	Nature 407:757-762(2000).				
DR	EMBL; AB002120; AAF30623.1;				
DR	GO; GO:0003824; P:catalytic activity; IEA.				
DR	InterPro; IPR01932; PP2C-like.				
DR	Pfam; PF00481; PP2C; 1.				
DR	SMART; SM00332; PP2C; 1.				
KW	SMART; SM00331; PP2C_SIC; 1.				
SO	SEQUENCE 247 AA; 28322 MW; AB1EP16AF275C80 CRC64;				
Query Match 91.1%; Score 41; DB 2; Length 247;					
Best Local Similarity 100.0%; Pred. No. 6.5;					
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 FLSTYDLFV 8				
DB	176 FLSTYDLFV 183				
RESULT 2					
ID	Q7BHL0	PRELIMINARY;	PRT;	69 AA.	
AC	Q7BHL0;				
DT	05-JUL-2004 (TREMBLrel. 27, Created)				
DT	05-JUL-2004 (TREMBLrel. 27, Last sequence update)				
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)				
DE	Opf (fragment).				
GN	Name=Opf;				
OS	Fusobacterium necrophorum.				
OC	Bacteria; Fusobacteriales; Fusobacteriaceae;				
OC	Fusobacterium.				
OX	NCBI_TaxID=859;				
RN	[1]				
RP	SEQUENCE FROM N.A.				

```

RC STRAIN=A25;
RA MEDLINE=21391803; PubMed=11500416;
RA Narayanan S.K., Nagaraaja T.G., Chengappa M.M., Stewart G.C.;
RT "Cloning, sequencing, and expression of the leukotoxin gene from
RT Fusobacterium necrophorum."
RL Infect. Immun. 69:5447-5455(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A25;
RA Oelke A.M., Nagaraaja T.G., Wilkerson M.J., Stewart G.C.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A25;
RA Stewart G.C., Narayanan S., Oelke A.M., Nagaraaja T.G., Chengappa M.M.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF312861; AAS45303.1; -.
PT NON TER 1
SQ SEQUENCE 69 AA; 8271 MW; 82443B7D96429965 CRC64;

Query Match
Best Local Similarity 84.4%; Score 38; DB 2; Length 69;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9
DB 16 FISHDLFV 24

RESULT 3
AAS45303 PRELIMINARY; PRT; 69 AA.
ID AAS45303;
AC AAS45303;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE Oppf (Fragment).
GN OPPF.
OS Fusobacterium necrophorum.
OC Bacteria; Fusobacteriia; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=859;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A25;
RA MEDLINE=21391803; PubMed=11500416;
RA Narayanan S.K., Nagaraaja T.G., Chengappa M.M., Stewart G.C.;
RT "Cloning, sequencing, and expression of the leukotoxin gene from
RT Fusobacterium necrophorum."
RL Infect. Immun. 69:5447-5455(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A25;
RA Oelke A.M., Nagaraaja T.G., Wilkerson M.J., Stewart G.C.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A25;
RA Stewart G.C., Narayanan S., Oelke A.M., Nagaraaja T.G., Chengappa M.M.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF312861; AAS45303.1; -.
PT NON TER 1
SQ SEQUENCE 69 AA; 8271 MW; 82443B7D96429965 CRC64;

Query Match
Best Local Similarity 84.4%; Score 38; DB 2; Length 69;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9
DB 16 FISHDLFV 24

```

```

RESULT 4
O99T82 PRELIMINARY; PRT; 259 AA.
ID O99T82;
AC O99T82;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Truncated transposase.
GN OrderedLocustNames=SAV1787;
OC Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158678;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M50 / ATCC 700699;
RA MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kohara S., Goto S., Yabuzaki J.,
RA Kanetsu M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hatori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003363; BAB57949.1; -.
DR PIR; C89964; C89964.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
DR Complete proteome.
SQ SEQUENCE 259 AA; 30921 MW; 73B1B908A649E4AC CRC64;

Query Match
Best Local Similarity 82.2%; Score 37; DB 2; Length 259;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FLSYDLF 7
DB 12 FLSYDLF 18

RESULT 5
O9X0X9 PRELIMINARY; PRT; 191 AA.
ID O9X0X9;
AC O9X0X9;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=TM1252;
OC Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RA MEDLINE=99287316; PubMed=10360571; DOI=10.1038/20601;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Hatt D.H., Hickey B.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L.A., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.L.,
RA Stewart A.M., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Heidelberg J.F., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
DR EMBL; AE001780; AAD36327.1; -.
DR PIR; F72277; F72277.
DR TIGR; TM1252; -.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 191 AA; 21566 MW; 82A6ABCD45D08498 CRC64;

```

Query Match 80.0%; Score 36; DB 2; Length 191;  
 Best Local Similarity 66.7%; Pred. No. 51;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLFV 9  
 |||||  
 DB 44 FLSDFPFI 52

RESULT 6

Q9D9A8 PRELIMINARY; PRT; 101 AA.  
 ID Q9D9A8  
 AC Q9D9A8  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1700113101 product:hypothetical protein, full insert sequence.  
 GN Name=1700113101Rik;  
 OS Mus musculus (Mouse);  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RA The FANTOM Consortium;  
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Kono H., Akiyama Y., Nishi K., Kitamura T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A., Yamamoto R., Maksumoto H., Sakaguchi S., Ikegami T., Kaibiwagi K., Fujikawa S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;

RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H., Arikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato F., Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numata R., Ono M., Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y., Toya T., Yamamoto T., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK007197; BAB24895.1;  
 DR MGD; MGI:1923914; 1700113101Rik.  
 DR GO; GO:0005085; F:guany1-nucleotide exchange factor activity; IEA.  
 DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.  
 DR InterPro; IPR000651; Rasgef\_N.  
 DR PROSITE; PS50212; Rasgef\_NTER; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 101 AA; 11635 MW; E9277CDA50D73D30 CRC64;

Query Match 77.8%; Score 35; DB 2; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8  
 |||||  
 DB 18 LSYDLFV 24

RESULT 7

Q9NP10 PRELIMINARY; PRT; 162 AA.  
 ID Q9NP10  
 AC Q9NP10  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE HSPC198 (Hypothetical protein HSPC198).  
 GN Name=HSPC198;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Umbilical cord blood;  
 RX MEDLINE=20499367; PubMed=11042152;  
 RA Zhang Q.H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G., Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.G., Zhang J.W., Tao J., Huang Q.H., Zhou J., Hu G.X., Gu J., Chen S.J., Chen Z.;  
 RT "Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells.";  
 RL Genome Res. 10:1546-1560(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bladder;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F., Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L., Stepien M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B., Brownstein M.J., Uscin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J., Bosak S.A., McEwan P.C., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Huliy S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton B., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,

RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Bladder;  
 RA Strauberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF151032; AAF36116.1; -;  
 DR EMBL; BC005201; AAH05201.1; -;  
 DR EMBL; AF151030; AAF36116.1; -;  
 KM Hypothetical protein.  
 SQ SEQUENCE 162 AA; 19262 MW; 31D76DD5C95333DE CRC64;  
 Query Match 77.8%; Score 35; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2 LSYDFV 8  
 Db 18 LSYDFV 24  
 RESULT 8  
 Q9D6G5 PRELIMINARY; PRT; 162 AA.  
 ID Q9D6G5  
 AC Q9D6G5;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)  
 DT 01-OCT-2004 (TrEMBLrel. 28, last annotation update)  
 DE Mus musculus adult male hippocampus cDNA, RIKEN full-length enriched  
 DE library, clone:2900055D14 product:hypothetical protein, full insert  
 DE sequence (Mus musculus adult pancreas islet cells cDNA, RIKEN full-  
 DE length enriched library, clone:C820001003 product:hypothetical  
 DE protein, full insert sequence) (RIKEN cDNA 2900055D14).  
 OS Name=2900055D14Rik;  
 GN Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Hippocampus, and Pancreas;  
 RC MEDLINE=99279253; PubMed=10349636;  
 RA Carininci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Hippocampus, and Pancreas;  
 RC MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Hippocampus, and Pancreas;  
 RC MEDLINE=20499374; PubMed=11042159;  
 RA Carininci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes";  
 RL Genome Res. 10:1617-1630(2000).

RN [5]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Hippocampus, and Pancreas;  
 RC MEDLINE=20530913; PubMed=11076861;  
 RX Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carininci P.,  
 RA Kono H., Akiyama U., Nishi K., Kitsuana T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada K.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasaiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system:384-format  
 RT sequencing pipeline with 384 multicapillary sequencer";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Hippocampus;  
 RC Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arakawa T., Bono H., Carininci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Pancreas;  
 RC Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carininci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kanakawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura Y.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohmoto N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Olfactory epithelium;  
 RC MEDLINE=22388257; PubMed=12477932;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,  
 RA Altschuler S.F., Zeeberg B., Bueltow R.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carininci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [9]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Olfactory epithelium;  
 RC Strauberg R.;  
 RT Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.



DR EMBL; AK013694; BAB28958.1; -  
 DR EMBL; AK050471; BAC34274.1; -  
 DR EMBL; BC058237; AAH58237.1; -  
 DR MGI; MGI:1920232; 290055D14R1k.  
 KM Hypothetical protein.  
 SQ SEQUENCE 162 AA; 19150 MW; 4E1D72AD014E29E0 CRC64;

Query Match 77.8%; Score 35; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
 |||||  
 Db 18 LSYDLFV 24

RESULT 9  
 AAH58237 PRELIMINARY; PRT; 162 AA.  
 AC AAH58237;  
 DT 02-MAR-2004 (TREMBlrel. 27, Created)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
 DE RIKEN CDNA 2900055D14.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Olfactory epithelium;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,  
 RA Hopkins R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
 RA Altschul S.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueclin T.B., Toshiyuki S., Carrinci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosnak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Krzywicki M.I., Skalek U., Smalins D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Olfactory epithelium;  
 RA Strausberg R.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC058237; AAH58237.1; -  
 SQ SEQUENCE 162 AA; 19150 MW; 4E1D72AD014E29E0 CRC64;

Query Match 77.8%; Score 35; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 69;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSYDLFV 8  
 |||||  
 Db 18 LSYDLFV 24

RESULT 10  
 O812F8 PRELIMINARY; PRT; 192 AA.  
 AC O812F8;

DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein PF11750c.  
 GN Name=PF11750c;  
 OS Plasmodium falciparum (isolate 3D7).  
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_Taxid=36329;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=22255708; PubMed=12368867;  
 RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,  
 RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,  
 RA Buckee C.O., Burrows C., Chevreach I., Chillingworth C.,  
 RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Cotton C.,  
 RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,  
 RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,  
 RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,  
 RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,  
 RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Leonard N.,  
 RA Lane A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,  
 RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitz E.,  
 RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,  
 RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,  
 RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;  
 RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";  
 RL Nature 419:527-531(2002).  
 DR EMBL; AL929359; CAD52036.1; -  
 DR InterPro; IPR006496; DUF\_Pfalc267.  
 DR TrEMBL; TIGR01609; PF\_unchar\_267; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 192 AA; 23573 MW; DCF5A5B8802C7A3 CRC64;

Query Match 77.8%; Score 35; DB 2; Length 192;  
 Best Local Similarity 55.6%; Pred. No. 82;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 9  
 |||||  
 Db 37 FLSYDMLII 45

RESULT 11  
 O84AF2 PRELIMINARY; PRT; 203 AA.  
 AC O84AF2;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Beryl.  
 GN Name=betyl;  
 OS Bacillus stearohermophilus.  
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.  
 OX NCBI\_Taxid=1422;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Y406;  
 RA Xu S.-Y., Samuelson J., Pelletier J., Sibley M., Wilson G.G.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY197779; AAC48714.1; -  
 SQ SEQUENCE 203 AA; 23193 MW; C6F5C22A35DB04EF CRC64;

Query Match 77.8%; Score 35; DB 2; Length 203;  
 Best Local Similarity 75.0%; Pred. No. 86;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLFV 8  
 |||||  
 Db 136 FVAYDLFV 143

RESULT 12

```

088216
ID Q88216 PRELIMINARY; PRT; 35 AA.
AC Q88216;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lipoprotein, putative.
OS OrderedLocustNames=PSPT02641;
OC Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RX MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA Bueli C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA Gwin M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Dougherty S.C., Brinkac L.M., Beaman M.J., Haft D.H.,
RA Nelson W.C., Davidesen T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA Kfour H.M., Pedorova N.B., Tran B., Russell D., Berry K.T.,
RA Utechtack T.R., Van Aken S.B., Feldblyum T.V., D'Ascenzo M.,
RA Deng W.-L., Ramos A.R., Alfano J.R., Carlinhour S., Chatterjee A.K.,
RA Belaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA Bender C.L., White O., Fraser C.M., Collier A.;
RT "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR EMBL; AE016865; AAO56144.1; -.
DR TIGR; PSPT02641; -.
KM Complete proteome; Lipoprotein.
SQ SEQUENCE 35 AA; 3854 MW; E209450422A75C09 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 35;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSTYDLFV 9
DB 9 FLSTYGLYVI 17

RESULT 13
0882D9 PRELIMINARY; PRT; 95 AA.
AC 0882D9; Q74VZ3; Q7CJW2; 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein YPO2589 (Hypothetical protein Y1159).
GN OrderedLocustNames=YPI126, YPO2589, Y1159;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parthill J., Wren B.W., Thomson N.R., Tildall R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holroyd S., Jagers K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,

```

```

RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Niles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM."
RL J. Bacteriol. 184:4601-4611(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ414152; CAC91389.1; -.
DR EMBL; AE013719; JAM84736.1; -.
DR EMBL; AE017131; AAS61372.1; -.
DR FIR; A10315; A10315.
DR InterPro; IPR007138; ABM.
DR InterPro; IPR011008; Dimer_A_B_barrel.
DR Pfam; PF03992; ABM; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 95 AA; 11144 MW; 13192FF15A3A9429 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 95;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8
DB 36 LSYDLFI 42

RESULT 14
AAS61372 PRELIMINARY; PRT; 95 AA.
AC AAS61372;
DT 24-MAR-2004 (TrEMBLrel. 27, Created)
DT 24-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN YPI126.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RX STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017131; AAS61372.1; -.
KM Hypothetical protein.
SQ SEQUENCE 95 AA; 11144 MW; 13192FF15A3A9429 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 95;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSYDLFV 8
DB 36 LSYDLFI 42

RESULT 15
061299 PRELIMINARY; PRT; 184 AA.
AC 061299;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

```



```

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
CC -1- CATALYTIC ACTIVITY: D-mannose + 2-dehydro-3-deoxy-D-gluconate +
CC H(2)O.
CC -1- PATHWAY: Hexuronate conversion to KDG.
CC -1- SIMILARITY: Belongs to the mannose dehydratase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP003010; BAB53075.1; -.
DR HAMAP; MF_00106; -.
DR InterPro; IPR004628; Man_dehyd.
DR Pfam; PF03786; Uxua; 1.
DR TIGRFAMs; TIGR00695; uxua; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 394 AA; 44542 MW; EBB0BC18FPC4003 CRC64;

Query Match 75.6%; Score 34; DB 1; Length 394;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLPV 9
DB 133 PAAVDFVL 141

RESULT 19
AC Q7VHG1 PRELIMINARY; PRT; 409 AA.
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ubiquinol cytochrome c oxidoreductase.
GN OrderedLocustNames=HH1006;
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51449 / 3B1;
RC MEDLINE=22709201; PubMed=12810954;
RA Suerbaum S., Josephans C., Sterzenbach T., Dreescher B., Brandt P.,
RA Bell M., Droege M., Partmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
DR EMBL; AB017147; AAP77603.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom B_C; 1.
DR Pfam; PF00033; Cytochrom B_N; 1.
DR PROSITE; PS00193; CYTOCHROME_B_OO; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 409 AA; 47066 MW; DEC0918DDCC055C49 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 409;

```

```

Best Local Similarity 77.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLPV 9
DB 252 FLSKDIFFV 260

RESULT 20
AC Q7MT01 PRELIMINARY; PRT; 411 AA.
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE UBIOJINOL CYTOCHROME C OXIDOREDUCTASE, CYTOCHROME B SUBUNIT
DE (PBCB)
GN Name=CITB, OrderedLocustNames=MS2153;
OS Wolinella succinogenes.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Wolinella.
OX NCBI_TaxID=844;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSMZ 1740;
RC MEDLINE=22882897; PubMed=14500908;
RA Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimek O.,
RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,
RA Meyer F., Lederer H., Schuster S.C.;
RT "Complete genome sequence and analysis of Wolinella succinogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).
DR EMBL; BX571662; CAE1147.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; Cytochrom B_C; 1.
DR Pfam; PF00033; Cytochrom B_N; 1.
DR PROSITE; PS00193; CYTOCHROME_B_OO; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 411 AA; 47252 MW; AS170467BBED61CF CRC64;

Query Match 75.6%; Score 34; DB 2; Length 411;
Best Local Similarity 77.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FLSYDLPV 9
DB 253 FLSKDIFFV 261

RESULT 21
AC Q26064 PRELIMINARY; PRT; 412 AA.
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ubiquinol cytochrome c oxidoreductase, cytochrome b subunit
DE (PBCB).
GN OrderedLocustNames=HP1539;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RC MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;
RA Tomb J.-F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S.R.,
RA Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F.,

```

RA Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,  
RA Glodok A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D.,  
RA Hickey E.K., Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D.,  
RA Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,  
RA Matthey L., Mallin E., Hayes W.S., Borodovsky M., Karp P.D.,  
RA Smith H.O., Fraser C.M., Venter J.C.,  
RT "The complete genome sequence of the gastric pathogen *Helicobacter*  
RT *pylori*.";  
RL Nature 388:539-547(1997).  
DR EMBL; AE000652; AAD08579.1; -.  
DR PIR; C64712; C64712.  
DR TIGR; HPI539; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR005798; Cytb b6 C.  
DR InterPro; IPR005797; Cytb b6 N.  
DR Pfam; PF00032; Cytochrom B\_C; 1.  
DR Pfam; PF00033; Cytochrom B\_N; 1.  
DR PROSITE; PS00193; CYTOCHROME\_B\_QQ; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 412 AA; 47510 MW; 954646D95A1F925A CRC64;

Query Match 75.6%; Score 34; DB 2; Length 412;  
Best Local Similarity 77.8%; Pred. No. 2.8e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLSTDLFV 9  
Db 253 FLSDIFV 261

RESULT 22  
09ZWT5 PRELIMINARY; PRT; 412 AA.

AC 09ZWT5  
DT 01-MAY-1999 (TRENBLREL. 10, Created)  
DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)  
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)  
DE Ubiquinol cytochrome c oxidoreductase, cytochrome b subunit.  
GN Name=petb, OrderedLocustNames=JHP1460.  
OS *Helicobacter pylori* J99 (Campylobacter *pylori* J99).  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
OC *Helicobacteraceae*; *Helicobacter*.  
OX NCBI\_Taxid=85963;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9912057; PubMed=9923682;  
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
RA Tummino P.J., Carno A., Uria-Nickelsen M., Mills D.M., Ives C.,  
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
RA Trust T.J.,  
RT "Genomic sequence comparison of two unrelated isolates of the human  
RT gastric pathogen *Helicobacter pylori*.";  
RL Nature 397:176-180(1999).  
DR EMBL; AE001568; AAD07046.1; -.  
DR PIR; D71803; D71803.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR005798; Cytb b6 C.  
DR InterPro; IPR005797; Cytb b6 N.  
DR Pfam; PF00032; Cytochrom B\_C; 1.  
DR Pfam; PF00033; Cytochrom B\_N; 1.  
DR PROSITE; PS00193; CYTOCHROME\_B\_QQ; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 412 AA; 47631 MW; 363F3CDB36380BD CRC64;

Query Match 75.6%; Score 34; DB 2; Length 412;  
Best Local Similarity 77.8%; Pred. No. 2.8e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FLSTDLFV 9  
Db 253 FLSDIFV 261

RESULT 23

09X019 PRELIMINARY; PRT; 452 AA.  
AC 09X019  
DT 01-NOV-1999 (TRENBLREL. 12, Created)  
DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)  
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)  
DE NAH dehydrogenase, putative.  
GN OrderedLocustNames=TM1105;  
OS Thermotoga maritima.  
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.  
OX NCBI\_Taxid=2336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=MSB8 / DSM 3109 / ATCC 43589;  
RX MEDLINE=99287316; PubMed=10360571; DOI=10.1038/20601;  
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,  
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
RA McDonald L.A., Utterback T.R., Nalek J.A., Linher K.D., Garrett M.M.,  
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.L.,  
RA Heidelberg J.F., Sutton G.G., Fleischmann R.D., Bisen J.A., White O.,  
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.,  
RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
RT genome sequence of *Thermotoga maritima*.";  
RL Nature 399:323-329(1999).  
CC -1-SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
DR EMBL; AE001769; AAD36181.1; -.  
DR PIR; C72295; C72295.  
DR TIGR; TM1105; -.

DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.  
DR InterPro; IPR001750; Oxidored q1.  
DR Pfam; PF00361; Oxidored q1; 1.  
KW Complete proteome; Transmembrane.  
SQ SEQUENCE 452 AA; 50538 MW; 19BEFCDE664F4B95 CRC64;

Query Match 75.6%; Score 34; DB 2; Length 452;  
Best Local Similarity 85.7%; Pred. No. 3.1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FLSTDLFV 7  
Db 102 FVSYDLF 108

RESULT 24

09ZWT3 PRELIMINARY; PRT; 516 AA.  
AC 09ZWT3  
DT 01-MAY-1999 (TRENBLREL. 10, Created)  
DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)  
DT 05-JUL-2004 (TRENBLREL. 27, Last annotation update)  
DE Inorganic phosphate transporter.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_Taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99087485; PubMed=9872450;  
RA Okumura S., Mitukawa N., Shirano Y., Shibata D.,  
RT "Phosphate transporter gene family of Arabidopsis thaliana.";  
RL DNA Res. 5:261-269(1996).  
RN [2]  
RP SEQUENCE FROM N.A.

Query Match 75.6%; Score 34; DB 2; Length 412;  
Best Local Similarity 77.8%; Pred. No. 2.8e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:31-63(2000).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AB005746; BAA33390.1; -.
DR EMBL; AB026638; BAA97413.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005315; F:inorganic phosphate transporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; PFS.
DR InterPro; IPR004738; Phos_pemasease.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR TIGRFAMs; TIGR00887; TA0109; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR Transmembrane.
SQ SEQUENCE 516 AA; 56246 MW; 97B2BFEFA894A145 CRC64;

Query Match
Best Local Similarity 75.6%; Score 34; DB 2; Length 516;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 3 SYDLFV 9
Db 37 SYDLFV 43

RESULT 25
Oy105 PRELIMINARY; PRT; 590 AA.
ID 08Y105;
AC 08Y105;
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE PROBABLE ACTIVATION/SECRETION PROTEIN.
GN Name=RS04850; OrderedLocustNames=RS00888;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OC NCBI_Taxid=305;
OX NCBI_Taxid=305;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mengent S.,
RA Ariat M., Billault A., Broctier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebaud P., Whalen M., Winkler P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AL646061; CAD14590.1; -.
DR InterPro; IPR005565; HlyB.
DR Pfam; PF03865; HlyB; 1.
DR Complete proteome.
SQ SEQUENCE 590 AA; 64535 MW; DC0D1907965D5A0D CRC64;

Query Match
Best Local Similarity 75.6%; Score 34; DB 2; Length 590;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LSYDLFV 8
Db 559 LSYDLFV 565

```

```

RESULT 26
Oy109 PRELIMINARY; PRT; 618 AA.
ID 07N109;
AC 07N109;
DT 01-MAR-2004 (T-EMBLrel. 26, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE NADH dehydrogenase subunit 5.
GN Name=ndhF; OrderedLocustNames=g112083;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Chroococcales; Gloeobacter.
OC NCBI_Taxid=33072;
OX NCBI_Taxid=33072;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Muro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Matsumoto M., Matsuno A., Nakazaki N., Shimo S.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids."
RL DNA Res. 10:137-145(2003).
DR EMBL; AP006575; BAC90024.1; -.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
DR InterPro; IPR003918; NADhub_oxred4.
DR InterPro; IPR003916; NADhub_oxred5.
DR InterPro; IPR010217; Ndh3_CO2.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR001516; Oxidored_q1_N.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF00662; Oxidored_q1_N; 1.
DR PRINTS; PR01434; NADHNDHANS5.
DR PRINTS; PR01437; NCOXDRDPTAS4.
DR TIGRFAMs; TIGR01960; ndhF_CO2; 1.
DR Complete proteome; NAD; NADP; Oxidoreductase; Plastocyanine; Quinone.
SQ SEQUENCE 618 AA; 66633 MW; D1EBFAB93BEFAD02 CRC64;

Query Match
Best Local Similarity 75.6%; Score 34; DB 2; Length 618;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FLSYDLFV 8
Db 534 FLSYDLFV 541

RESULT 27
KBT6 HUMAN STANDARD; PRT; 674 AA.
ID KBT6 HUMAN
AC 08Y67; O8NBL0; O8NDMS; O96MP6;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Kelch repeat and BTB domain containing protein 6.
GN Name=KBTBD6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_Taxid=9606;
OX NCBI_Taxid=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Magatsuna M.,
RA Shitatori A., Sudo H., Hosoliti T., Kaku Y., Kodaira H., Kondo H.,

```

RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,  
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,  
 RA Yamazaki M., Nimomiya K., Ishibashi T., Yamashita H., Murakawa K.,  
 RA Fujimori K., Tanai H., Kimata H., Hironaka S., Chiba Y.,  
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Horita T.,  
 RA Kusano J., Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O.,  
 RA Nomura Y., Togaiya S., Komai F., Hara R., Takeuchi K., Arita M.,  
 RA Imobe N., Mueasano K., Yuuki H., Oshima A., Sasaki N., Aotaka S.,  
 RA Yoshioka Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi T.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
 RA Kawabata A., Hikiji T., Kobatake S., Inagaki H., Ikema Y., Okamoto S.,  
 RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigetani K., Senda T.,  
 RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.";  
 RL Nat. Genet. 36:40-45(2004).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain, and Prostate;  
 RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datschenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uscin T.B., Toshylyuk S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fehey J., Holton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]

RP SEQUENCE OF 21-674 FROM N.A.  
 RC TISSUE=Brain;  
 RA Pousetka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.,  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 B7B/POZ domain.  
 CC -1- SIMILARITY: Contains 6 Kelch repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

CC  
 DR EMBL: AK056633; BAB71238.1; -  
 DR EMBL: AK096608; BAC04826.1; ALT\_INIT.  
 DR EMBL: BC000560; AA000560.1; -  
 DR EMBL: BC051349; AA051349.1; -  
 DR EMBL: AL831839; CAD38699.1; -  
 DR Genev, HGNC:25340; KERTD6.  
 DR InterPro: IPR000210; B7B\_POZ.  
 DR InterPro: IPR011043; Gal\_oxid\_central.

DR InterPro: IPR006652; Kelch\_rep.  
 DR Pfam: PF00651; B7B; 1.  
 DR Pfam: PF01344; Kelch; 1.  
 DR PROSITE: PS0097; B7B; 1.

KW Kelch repeat; Repeat.  
 FT DOMAIN 63 138  
 FT REPEAT 386 435  
 FT REPEAT 436 484  
 FT REPEAT 486 523  
 FT REPEAT 524 564  
 FT REPEAT 567 616  
 FT REPEAT 642 673  
 FT CONFLICT 61  
 SQ SEQUENCE 674 AA; 76137 MW; SED49AABC6350CB5 CRC64;  
 L -> Q (in Ref. 1; BAB71238).

Query Match 75.6%; Score 34; DB 1; Length 674;  
 Best Local Similarity 66.7%; Pred. No. 4.6e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSTYDLFV 9  
 Db 474 FLSTYDLFV 482

RESULT 28  
 ID 091L83 PRELIMINARY; PRT; 729 AA.  
 AC 091L83;

DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE ORF169.  
 OS White spot syndrome virus (WSSV).  
 OC Viruses; dsDNA viruses; no RNA stage; Nimaviridae; Whispovirus.  
 CX NCBI\_TaxID=92652;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21342572; PubMed=11448154;  
 RA van Hulten M.C.W., Wittevelde J., Peters S., Kloosterboer N.,  
 RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.,  
 RT "The white spot syndrome virus DNA genome sequence.";  
 RL Virology 286:7-22(2001).  
 DR EMBL: AF369029; AAK77838.1; -

SQ SEQUENCE 729 AA; 85128 MW; 91816B756F02B3A CRC64;

Query Match 75.6%; Score 34; DB 2; Length 729;  
 Best Local Similarity 55.6%; Pred. No. 4.9e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FLSTYDLFV 9  
 Db 705 FLSTYDLFV 713

RESULT 29  
 ID 06CFR0 PRELIMINARY; PRT; 924 AA.  
 AC 06CFR0;

DT 01-OCT-2004 (TREMBlrel. 28, Created)  
 DT 01-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Similarities with tr|Q12527 Saccharomyces cerevisiae YPR049c related  
 DE to UG01P  
 GN ORFNames=YAL10B045989;

OS Yarrowia lipolytica (Candida lipolytica).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Dipodascaceae; Yarrowia.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CL1899;  
 RG GENOLEVURES;  
 RA Dujon B., Sherman D., Fischer G., Durrens P., Caaregola S.,

RA Lafontaine I., de Montigny J., Marck C., Neveu-Jisse C., Talla E.,  
 RA Goffard N., Frangoul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
 RA Barney S., Bianchini S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boismere A., Boyer J., Catolico L., Confanioli F., de Darivar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dunazet H., Gropi A.,  
 RA Hantay F., Hennequin C., Jaumaux N., Joyet P., Kachouri R.,  
 RA Kerrest A., Koszul R., Lemaitre M., Leduc I., Ma L., Muller H.,  
 RA Nicard J.M., Nikolski M., Oztas S., Olier-Kalogeropoulos O.,  
 RA Pollenz S., Potier S., Richard G.F., Strub M.L., Suleau A.,  
 RA Swemene D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zenon-Meyer M., Zivanovic I., Boloctin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Cautron B., Scarpell C., Gaillardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts."  
 RL Nature 430:35-44(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CLIB99;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR382128; CAG8729.1; -;  
 SQ SEQUENCE 924 AA; 104020 MW; C1E08535E5CB5AA03 CRC64;  
 Query Match 75.6%; Score 34; DB 2; Length 924;  
 Best Local Similarity 75.0%; Pred. No. 6.3e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLSYDLFV 8  
 Db 786 FLDYDLV 793  
 RESULT 30  
 ID Q8IDC3 PRELIMINARY; PRT; 1559 AA.  
 AC Q8IDC3;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein Phat60.  
 GN Name=Phat60; Synonyms=MAL13P1.286, PFI3\_0311, Phat61;  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
 OX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,  
 RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,  
 RA Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrett B.,  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL844509; CAD52702.1; -;  
 KM Hypothetical protein.  
 SQ SEQUENCE 1559 AA; 187166 MW; E7F2B3BE0E8B17 CRC64;  
 Query Match 75.6%; Score 34; DB 2; Length 1559;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+03;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLSYDLFV 8  
 Db 1080 FLSYSLFI 1087  
 RESULT 31  
 ID NU3M ARTSA STANDARD; PRT; 70 AA.  
 AC P19043;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase chain 3 (EC 1.6.5.3) (Fragments).  
 GN Name=ND3;  
 OS Artemia salina (Brine shrimp).

OG Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;  
 OC Artemiidae; Artemia.  
 OX NCBI\_TaxID=85549;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88289417; PubMed=3135541;  
 RA Batueca B., Garesse R., Calleja M., Valverde J.R., Marco R.;  
 RT "Genome organization of Artemia mitochondrial DNA."  
 RL Nucleic Acids Res. 16:6515-6529(1988).  
 CC -1-CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC -1-SIMILARITY: Belongs to the complex I subunit 3 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X07664; CAA30511.1; -;  
 DR EMBL; X07665; CAA30512.1; -;  
 DR InterPro; IPR000440; Oxidored\_g4.  
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.  
 FT NON\_TER 1  
 FT NON\_CONS 31 32  
 FT NON\_TER 70 70  
 SQ SEQUENCE 70 AA; 8443 MW; C2D7A01B98AF62D0 CRC64;  
 Query Match 73.3%; Score 33; DB 1; Length 70;  
 Best Local Similarity 75.0%; Pred. No. 76;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLSYDLFV 8  
 Db 33 FLSYPLFI 40  
 RESULT 32  
 ID Q45405 PRELIMINARY; PRT; 203 AA.  
 AC Q45405;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Spag.  
 GN Name=spag;  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC6633;  
 RX MEDLINE=93175869; PubMed=8439156;  
 RA Klein C., Kaletka C., Entian K.D.;  
 RT "Biosynthesis of the antibiotic subtilin is regulated by a histidine  
 RT kinase/response regulator system."  
 RL Appl. Environ. Microbiol. 59:296-303(1993).  
 DR PIR; I40517; I40517.1; -;  
 SQ SEQUENCE 203 AA; 22324 MW; F5045CAE34DCA347 CRC64;  
 Query Match 73.3%; Score 33; DB 2; Length 203;  
 Best Local Similarity 72.5%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LSYDLFV 9



Db 78 LSYDIFM 85

## RESULT 33

09CEI4 PRELIMINARY; PRT; 305 AA.  
AC 09CEI4;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
DE Oxidoreductase.  
GN Name=yj8; OrderedLocustNames=IL1823;  
OC Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OX NCBI\_TaxID=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IL1403;  
RX MEDLINE=21235186; PubMed=11337471;  
RA Boletín A., Wincker P., Manger S., Jallón O., Malarme K.,  
RA Weissenbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
RT lactis ssp. lactis IL1403."  
RL Genome Res. 11:731-753(2001).  
DR EMBL; AE006412; AAK05921.1; -.  
DR PIR; G86852; G86852.  
DR GO; GO:0016491; P:oxidoreductase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000683; GFO\_IDH/MOCA\_N.  
DR Pfam; PF01408; GFO\_IDH\_MOCA\_1.  
KW Complete proteome.  
SQ SEQUENCE 305 AA; 34663 MW; 29E857F0D936FC0B CRC64;

Query Match 73.3%; Score 33; DB 2; Length 305;  
Best Local Similarity 62.5%; Pred. No. 3.3e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 FLSYDLFV 8  
Db 155 FLMDYDIF 162

RESULT 34  
072HT6 PRELIMINARY; PRT; 332 AA.  
AC 072HT6;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames=TTCl400;  
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).  
OC Bacteria; Deinococcus-Thermus; Thermococci; Thermales; Thermaceae;  
OC Thermus.  
OX NCBI\_TaxID=262724;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC PubMed=15064768;  
RX Henne A., Brueggemann H., Raasch C., Wierzer A., Hartech T.,  
RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Atlas R.,  
RA Jacobl C., Starkvliene V., Schlenczek S., Dencker S., Huber R.,  
RA Klank H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;  
RT "The genome sequence of the extreme thermophile Thermus  
RT thermophilus."  
RL Nat. Biotechnol. 22:547-553(2004).  
DR EMBL; AE017305; AAS81742.1; -.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 332 AA; 36790 MW; 9192C1AB6872E5B CRC64;

Query Match 73.3%; Score 33; DB 2; Length 332;  
Best Local Similarity 85.7%; Pred. No. 3.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FLSYDLF 7  
Db 158 FLSFDLF 164

## RESULT 35

AAS81742 PRELIMINARY; PRT; 332 AA.  
AC AAS81742;  
DT 14-APR-2004 (TREMBlrel. 27, Created)  
DT 14-APR-2004 (TREMBlrel. 27, Last sequence update)  
DT 11-MAY-2004 (TREMBlrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN TTCl400.  
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).  
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;  
OC Thermus.  
OX NCBI\_TaxID=262724;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC PubMed=15064768;  
RX Henne A., Brueggemann H., Raasch C., Wierzer A., Hartech T.,  
RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Atlas R.,  
RA Jacobl C., Starkvliene V., Schlenczek S., Dencker S., Huber R.,  
RA Klank H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;  
RT "The genome sequence of the extreme thermophile Thermus  
RT thermophilus."  
RL Nat. Biotechnol. 22:547-553(2004).  
DR EMBL; AE017305; AAS81742.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 332 AA; 36790 MW; 9192C1AB6872E5B CRC64;

Query Match 73.3%; Score 33; DB 2; Length 332;  
Best Local Similarity 85.7%; Pred. No. 3.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 FLSYDLF 7  
Db 158 FLSFDLF 164

RESULT 36  
08RD34 PRELIMINARY; PRT; 355 AA.  
AC 08RD34;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Beta-lactamase class C and other penicillin binding proteins.  
GN Name=ampC; OrderedLocustNames=TTB0213;  
OS Thermobacterium tengcongensis.  
OC Bacteria; Firmicutes; Clostridia; Thermobacteriales;  
OC Thermobacteriaceae; Clostridiaceae; Thermobacter.  
OX NCBI\_TaxID=119072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MB4;  
RX MEDLINE=21992816; PubMed=11997336;  
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
RT "A complete sequence of the T. tengcongensis genome."  
RL Genome Res. 12:689-700(2002).  
DR EMBL; AE012994; AAM23514.1; -.  
DR InterPro; IPR001466; Beta\_lactamase.  
DR Pfam; PF00144; Beta\_lactamase; 1.  
KW Complete proteome.  
SQ SEQUENCE 355 AA; 40303 MW; 3FD7B0AD4AE46545 CRC64;

Query Match 73.3%; Score 33; DB 2; Length 355;  
Best Local Similarity 75.0%; Pred. No. 3.8e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FLSTYDLFV 8  
DB 257 FLSTYDLFV 264

RESULT 37  
MPK3\_ARATH STANDARD; PRT; 370 AA.  
ID MPK3\_ARATH  
AC Q39023; QM1E3; 35, Created)  
DT 01-NOV-1997 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Mitogen-activated protein kinase homolog 3 (EC 2.7.1.37) (MAP kinase 3) (AtMPK3).  
GN Name=MPK3; OrderedLocNames=At3g45640; ORFNames=T69.4, P9X21.220;  
OS Arabidopsis thaliana (Mouse-ear cress). Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Eudicotyledons; core eudicots; rosids;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=94109583; PubMed=8282107;  
RA Mizoguchi T., Hayashida N., Yamaguchi-Shinozaki K., Kamada H.,  
RA Shinozaki K.;  
RA "ATMPK3: a gene family of plant MAP kinases in Arabidopsis thaliana";  
RL PEST Lett. 336:440-444 (1993).  
[2]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=21016720; PubMed=1130713; DOI=10.1038/35048706;  
RA Salomondat M., Lemcke K., Rieger M., Ansoorge W., Unseld M.,  
RA Farmanan B., Valle G., Bioecker H., Perez-Alonso M., Obermaier B.,  
RA Delbeny M., Boutry M., Grivell L.A., Macho R., Pulgomech P.,  
RA De Simone V., Choise N., Artiguenave F., Robert C., Brothier P.,  
RA Wincker P., Catolico L., Weissenbach J., Sautin W., Quetier F.,  
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
RA Wurnbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,  
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nakamura G.,  
RA Verzi A., D'Angelo M., Pallavicini A., Toppi S., Simionati B.,  
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordisik G.,  
RA Reichelt J., Schaefer M., Schoen O., Bargues M., Terol J., Climent J.,  
RA Navarro P., Collado C., Perez-Perez A., Ottensmeyer B., Duchemin D.,  
RA Cooke R., Lande M., Berger-lano C., Punnelle B., Maury D.,  
RA de Haan W., Maier A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,  
RA Montfort A., Argitrou A., Flores M., Liguori R., Vitale D.,  
RA Manthaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
RA Rooney T., Rizzo M., Wates A., Uteback T., Fujii C.Y., Shea T.P.,  
RA Cressey T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,  
RA Pal G., Miltschker J., Sellers P., Gill J.E., Feldblyum T.V.,  
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,  
RA Frazer C.M., Kaneo T., Nakamura Y., Sato S., Kato T., Asanuma E.,  
RA Sasamoto S., Kimura T., Ideasa K., Kawashima K., Kishida Y.,  
RA Kiyokawa C., Kohara N., Matsuno M., Matsuno A., Muraki A.,  
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,  
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;  
RA "Sequence and analysis of chromosome 3 of the plant Arabidopsis thaliana";  
RL Nature 408:820-822 (2000).  
[3]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;  
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
RA Arakawa T., Banh U., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjani M., Hansen N.F.,

RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,  
RA Khan S., Koesema B., Ishida J., Jiang P.X., Jones T., Kawai J.,  
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
RA Satou M., Tame R., Yabeberg M., Wallenberg B.K., Wong C., Yamamura Y.,  
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
RT "Empirical analysis of transcriptional activity in the Arabidopsis genome";  
RL Science 302:842-846 (2003).  
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -1- COFACTOR: Magnesium (By similarity).  
CC -1- ENZYME REGULATION: Activated by cytosine and threonine phosphorylation (By similarity).  
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP kinase subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D21839; BAA04866.1; --  
CC EMBL; AL138657; CAB75493.1; --  
CC EMBL; AL157735; -- NOT ANNOTATED\_CDS.  
CC EMBL; AF386961; AAK62406.1; --  
CC PIR; S40459; S40459.  
CC PIR; T47504; T47504.  
CC HSSP; Q16538; 1KVI.  
CC Inactive; Q39023; --  
CC GeneFarm; 828; 89.  
CC InterPro; IPR008351; JNK\_MAPK.  
CC InterPro; IPR011009; Kinase like.  
CC InterPro; IPR003527; MAP kin.  
CC InterPro; IPR008352; p38-MAPK.  
CC InterPro; IPR007719; Prot\_Kinase.  
CC InterPro; IPR008271; Ser\_Thr\_pkin\_AS.  
CC Pfam; PF00069; Pkinase; 1.  
CC PRINTS; PR01772; JNKMAPKINASE.  
CC PRINTS; PR01773; p38MAPKINASE.  
CC ProDom; PD000001; Prot\_kinase; 1.  
CC PROSITE; PS01351; MAPK; 1.  
CC PROSITE; PS01017; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
CC PROSITE; PS50108; PROTEIN\_KINASE\_ST; 1.  
CC Kyr ATP-binding; Multigene family; Phosphorylation; Serine/threonine-protein kinase; Transferase.  
KW DOMAIN 38 324 Protein kinase.  
FT NP\_BIND 44 52 ATP (By similarity).  
FT BINDING 67 67 ATP (By similarity).  
FT ACT\_SITE 164 164 Proton acceptor (By similarity).  
FT MOD\_RES 196 196 Phosphothreonine (activates the kinase) (By similarity).  
FT MOD\_RES 198 198 Phosphotyrosine (activates the kinase) (By similarity).  
FT FT E -> D (in Ref. 1).  
SQ SEQUENCE 370 AA; 42716 MW; 69922AD79F3C9841 CRC64;  
Query Match 73.3%; Score 33; DB 1; Length 370;  
Best Local Similarity 71.4%; Pred. No. 4e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
OY 1 FLSTYDLFV 7  
DB 21 FLSTYDLFV 27

RESULT 38  
AA15326 PRELIMINARY; PRT; 370 AA.  
ID AA15326  
AC AA15326  
DT 02-MAR-2004 (TrEMBLrel. 27, Created)

DT 02-MAR-2004 (TReMBLrel. 27, last sequence update)  
 DE Mitogen-activated protein kinase 3.  
 GN ATG454640.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,  
 RA Palm C.J., Bowser L., Jones I., Bahn J., Carninci P., Chen H.,  
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,  
 RA Eckert J., Theologis A., Davis R.W.;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL, BT000007; AAI15326.1; -.  
 KW Kinase.  
 SQ SEQUENCE 370 AA; 42716 MW; 6992A7D97F3C9841 CRC64;  
 QY 1 FLSYDLFV 7  
 Db 21 FLSYDLFV 27  
 Query Match 73.3%; Score 33; DB 2; Length 370;  
 Best Local Similarity 71.4%; Pred. No. 4e+02; Indels 0; Gaps 0;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 RESULT 39  
 Q84542 PRELIMINARY; PRT; 432 AA.  
 ID 084542  
 AC 084542;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)  
 DE Similar to Aquifex cellulose synthase, corresponds to GenBank  
 DE Accession Number AE000738.  
 GN Name=A222/226R;  
 OS Paramecium bursaria chlorella virus 1 (PBCV-1).  
 OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.  
 OX NCBI\_TaxID=10506;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=9618795; Pubmed=8614977;  
 RA Lu Z., Li Y., Que Q., Kutish G.F., Rock D.L., Van Etten J.L.;  
 RT "Analysis of 94 kb of the chlorella virus PBCV-1 330-kb genome: map  
 RT positions 88 to 182."  
 RL Virology 216:102-123(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20013326; Pubmed=10544099;  
 RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,  
 RA Lasec A.D., Nickerson K.W., Van Etten J.L.;  
 RT "Chlorella virus PBCV-1 encodes a functional homospervimidine  
 RT synthase."  
 RL Virology 263:254-262(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20478054; Pubmed=11021991;  
 RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;  
 RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus  
 RT PBCV-1."  
 RL Virology 276:27-36(2000).  
 DR EMBL, U42580; AAC96590.2; -.  
 DR PIR, T17712; T17712.  
 DR PIR, T17717; T17717.  
 SQ SEQUENCE 432 AA; 49679 MW; 6A995363CF2B8CF CRC64;  
 Query Match 73.3%; Score 33; DB 2; Length 432;  
 Best Local Similarity 75.0%; Pred. No. 4.7e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 FLSYDLFV 8  
 Db 254 FLSYDLFV 261  
 Query Match 73.3%; Score 33; DB 2; Length 479;  
 Best Local Similarity 44.4%; Pred. No. 5.2e+02;  
 Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;  
 RESULT 40  
 Q9DHK3 PRELIMINARY; PRT; 479 AA.  
 ID 09DHK3  
 AC 09DHK3;  
 DT 01-MAR-2001 (TReMBLrel. 16, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)  
 DE 110R protein.  
 GN Name=110R;  
 OS Yaba-like disease virus (YLDV).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Yatapoxvirus.  
 OX NCBI\_TaxID=132475;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21176365; Pubmed=11277691;  
 RA Lee H.J., Essani K., Smith G.L.;  
 RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus."  
 RL Virology 281:170-192(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lee H.J.;  
 RL Thesis (2000), Sir William Dunn School of Pathology, University of  
 RL Oxford, Oxford, UNITED KINGDOM.  
 DR EMBL, AJ293568; CAC21348.1; -.  
 DR GO, GO:0005524; F:ATP binding; IEA.  
 DR GO, GO:0008026; F:ATP-dependent helicase activity; IEA.  
 DR GO, GO:0003676; F:nucleic acid binding; IEA.  
 DR GO, GO:0015668; F:type III site-specific deoxyribonuclease ac. .; IEA.  
 DR GO, GO:0009307; P:DNA restriction; IEA.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR InterPro: IPR006935; ResIII.  
 DR Pfam: PF00271; Helicase\_C; 1.  
 DR Pfam: PF04851; ResIII; 1.  
 DR SMART, SM00487; DEXDC; 1.  
 SQ SEQUENCE 479 AA; 55646 MW; 3057930013EBAA67 CRC64;  
 Query Match 73.3%; Score 33; DB 2; Length 479;  
 Best Local Similarity 44.4%; Pred. No. 5.2e+02;  
 Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FLSYDLFV 9  
 Db 196 YLNDYDFIL 204

Search completed: January 12, 2005, 20:14:13  
 Job time : 89 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:46:44 ; Search time 21.6 Seconds

(without alignments)  
27.633 Million cell updates/sec

Title: US-09-870-216C-11

Perfect score: 43

Sequence: 1 NLOQLMDRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PC/US.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	81.4	113	4	US-09-621-976-5972
2	35	81.4	152	4	US-09-513-999C-5230
3	33	76.7	90	4	US-09-248-796A-20334
4	32	74.4	283	4	US-09-583-110-3146
5	31	72.1	152	4	US-09-583-110-5027
6	31	72.1	159	4	US-09-248-796A-27510
7	31	72.1	189	4	US-09-248-796A-19441
8	31	72.1	270	4	US-09-602-787A-676
9	31	72.1	276	4	US-09-252-991A-18006
10	31	72.1	542	4	US-09-198-452A-496
11	31	72.1	640	4	US-09-248-796A-16129
12	31	72.1	855	4	US-09-248-796A-18667
13	30	69.8	67	4	US-09-252-991A-25165
14	30	69.8	168	4	US-09-270-767-33712
15	30	69.8	168	4	US-09-270-767-48929
16	30	69.8	231	4	US-09-540-236-2077
17	30	69.8	412	4	US-09-134-000C-6666
18	30	69.8	656	4	US-09-252-991A-29575
19	30	69.8	673	3	US-09-134-001C-4480
20	30	69.8	808	4	US-09-489-039A-11547
21	30	69.8	1116	4	US-09-583-110-5094
22	30	69.8	1784	3	US-09-040-738-2
23	30	69.8	1784	3	US-08-652-426A-2
24	29	67.4	133	4	US-09-270-767-38268
25	29	67.4	133	4	US-09-270-767-53485
26	29	67.4	228	3	US-09-134-001C-4694
27	29	67.4	228	3	US-09-134-001C-5495

28	29	67.4	269	4	US-09-818-780-93	Sequence 93, Appl
29	29	67.4	280	4	US-09-252-991A-30349	Sequence 30349, A
30	29	67.4	340	4	US-10-067-443-6	Sequence 6, Appl1
31	29	67.4	369	4	US-09-252-991A-23698	Sequence 23698, A
32	29	67.4	420	4	US-09-252-991A-30109	Sequence 30109, A
33	29	67.4	508	4	US-09-248-796A-15391	Sequence 15391, A
34	29	67.4	509	4	US-09-252-991A-32687	Sequence 22687, A
35	29	67.4	573	4	US-09-252-991A-18744	Sequence 18744, A
36	29	67.4	575	4	US-09-252-991A-29441	Sequence 29441, A
37	29	67.4	587	2	US-08-871-266B-18	Sequence 18, Appl
38	29	67.4	587	2	US-08-871-266B-18	Sequence 18, Appl
39	29	67.4	587	3	US-08-871-267B-24	Sequence 24, Appl
40	29	67.4	587	3	US-09-618-419-24	Sequence 24, Appl
41	29	67.4	856	4	US-09-252-991A-21444	Sequence 21444, A
42	29	67.4	1216	4	US-09-583-110-3824	Sequence 1824, Ap
43	29	67.4	1757	4	US-09-724-126A-15	Sequence 15, Appl
44	28	65.1	118	4	US-09-252-991A-27139	Sequence 27139, A
45	28	65.1	135	2	US-08-383-621-5	Sequence 5, Appl1

#### ALIGNMENTS

```

RESULT 1
US-09-621-976-5972
; Sequence 5972, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5972
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 75
; OTHER INFORMATION: Xaa = *, Glu, Lys, Gln
US-09-621-976-5972

Query Match      81.4%; Score 35; DB 4; Length 113;
Best Local Similarity 87.5%; Pred. No. 5;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy      1 NLOQLMDR 8
Db      23 NLOQLMDR 30

RESULT 2
US-09-513-999C-5230
; Sequence 5230, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59. US2. REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm

```

SEQ ID NO 5230  
LENGTH: 152  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 80  
OTHER INFORMATION: Xaa=Lys or Gln or Arg  
US-09-513-999C-5230

Query Match 81.4%; Score 35; DB 4; Length 152;  
Best Local Similarity 87.5%; Pred. No. 7;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOQLMDR 8  
DB 23 NLOQLVDR 30

RESULT 3  
US-09-248-796A-20334  
Sequence 20334, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstein et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248, 796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074, 725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096, 409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 20334  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-20334

Query Match 76.7%; Score 33; DB 4; Length 90;  
Best Local Similarity 66.7%; Pred. No. 10;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLOQLMDR 9  
DB 74 NLOQLFMDHI 82

RESULT 4  
US-09-583-110-3146  
Sequence 3146, Application US/09583110  
Patent No. 6699703  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
FILE REFERENCE: PATH00-07A  
CURRENT APPLICATION NUMBER: US/09/583, 110  
CURRENT FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/107, 433  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/085, 131  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: US 60/051, 553  
PRIOR FILING DATE: 1997-07-02  
NUMBER OF SEQ ID NOS: 5322  
SEQ ID NO 3146  
LENGTH: 283  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-583-110-3146

Query Match 74.4%; Score 32; DB 4; Length 283;  
Best Local Similarity 66.7%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 9  
DB 172 NLOQLIDRI 180

RESULT 5  
US-09-583-110-5027  
Sequence 5027, Application US/09583110  
Patent No. 6699703  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
FILE REFERENCE: PATH00-07A  
CURRENT APPLICATION NUMBER: US/09/583, 110  
CURRENT FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/107, 433  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/085, 131  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: US 60/051, 553  
PRIOR FILING DATE: 1997-07-02  
NUMBER OF SEQ ID NOS: 5322  
SEQ ID NO 5027  
LENGTH: 152  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-583-110-5027

Query Match 72.1%; Score 31; DB 4; Length 152;  
Best Local Similarity 55.6%; Pred. No. 47;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 9  
DB 132 NLOQLISRI 140

RESULT 6  
US-09-248-796A-27510  
Sequence 27510, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstein et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248, 796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074, 725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096, 409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 27510  
LENGTH: 159  
TYPE: PRT  
ORGANISM: Candida albicans  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (151)  
OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown  
US-09-248-796A-27510

Query Match 72.1%; Score 31; DB 4; Length 159;  
Best Local Similarity 66.7%; Pred. No. 50;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLOLMDRV 9  
|||:|:|  
Db 19 NLOIKKQV 27

RESULT 7  
US-09-248-796A-19441  
Sequence 19441, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Mainstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
FILE REFERENCE: 107196.132  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 19441  
LENGTH: 189  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-19441

Query Match 72.1%; Score 31; DB 4; Length 189;  
Best Local Similarity 66.7%; Pred. No. 60;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLOLMDRV 9  
|||:|:|  
Db 28 NLOIKKQV 36

RESULT 8  
US-09-602-787A-676  
Sequence 676, Application US/09602787A  
Patent No. 6696561  
GENERAL INFORMATION:  
APPLICANT: Pompeius, Mark  
APPLICANT: Krüger, Burkhard  
APPLICANT: Schöder, Hartwig  
APPLICANT: Zelder, Oskar  
APPLICANT: Haberhauer, Gregor  
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE  
FILE REFERENCE: BGI-125CP  
CURRENT APPLICATION NUMBER: US/09/602,787A  
CURRENT FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: US/09/602,787A  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: DE 19931454.3  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931478.0  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19931563.9  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19932122.1  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932124.8  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932125.6  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932128.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932180.9  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932182.5  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932190.6

PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932191.4  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932209.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932212.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932227.9  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932228.7  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932229.5  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932230.9  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932927.3  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933005.0  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933006.9  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19940764.9  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940765.7  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940766.5  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940830.0  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940831.9  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940832.7  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19940833.5  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: DE 19941378.9  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941379.7  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941395.9  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19942077.7  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942078.5  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942079.3  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: DE 19942088.2  
PRIOR FILING DATE: 1999-09-03  
NUMBER OF SEQ ID NOS: 676  
SEQ ID NO 676  
LENGTH: 270  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-602-787A-676

Query Match 72.1%; Score 31; DB 4; Length 270;  
Best Local Similarity 66.7%; Pred. No. 89;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLOLMDRV 9  
|||:|:|  
Db 22 NLOLMDRV 30

RESULT 9  
US-09-252-991A-18006  
Sequence 18006, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 18006  
LENGTH: 276  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (53)  
OTHER INFORMATION: Identity of amino acid at the above locations are unknown.  
US-09-252-991A-18006

Query Match  
Best Local Similarity 72.1%; Score 31; DB 4; Length 276;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9  
||:|||||  
Db 60 HLELVLDV 68

RESULT 10  
US-09-198-452A-496  
Sequence 496, Application US/09198452A  
Patent No. 6559294  
GENERAL INFORMATION:  
APPLICANT: Griffiths, R.  
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
FILE REFERENCE: 9710-003-999  
CURRENT APPLICATION NUMBER: US/09/198,452A  
CURRENT FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 6849  
SEQ ID NO 496  
LENGTH: 542  
TYPE: PRT  
ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-496

Query Match  
Best Local Similarity 72.1%; Score 31; DB 4; Length 542;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9  
||:|||||  
Db 531 NLOQLASDRI 539

RESULT 11  
US-09-248-796A-16129  
Sequence 16129, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstein et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 16129  
LENGTH: 640

TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-16129

Query Match  
Best Local Similarity 72.1%; Score 31; DB 4; Length 640;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9  
||:|||||  
Db 12 NKLILDKI 20

RESULT 12  
US-09-248-796A-18667  
Sequence 18667, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstein et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 18667  
LENGTH: 855  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-18667

Query Match  
Best Local Similarity 72.1%; Score 31; DB 4; Length 855;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9  
||:|||||  
Db 493 NLRLLDHV 501

RESULT 13  
US-09-252-991A-25165  
Sequence 25165, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 25165  
LENGTH: 67  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25165

Query Match  
Best Local Similarity 69.8%; Score 30; DB 4; Length 67;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QLVMDRV 9  
||:|||||  
Db 47 QLVMDRV 53



RESULT 14  
US-09-270-767-33712  
Sequence 33712, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 33712  
LENGTH: 168  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-270-767-33712

Query Match 69.8%; Score 30; DB 4; Length 168;  
Best Local Similarity 62.5%; Pred. No. 85;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQIMDRV 9  
||:||||:  
DB 84 LELLDRI 91

RESULT 15  
US-09-270-767-48929  
Sequence 48929, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 48929  
LENGTH: 168  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-270-767-48929

Query Match 69.8%; Score 30; DB 4; Length 168;  
Best Local Similarity 62.5%; Pred. No. 85;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQIMDRV 9  
||:||||:  
DB 84 LELLDRI 91

RESULT 16  
US-09-540-236-2077  
Sequence 2077, Application US/09540236  
Patent No. 6673910  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540,236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840  
SEQ ID NO 2077  
LENGTH: 231  
TYPE: PRT  
ORGANISM: M.catarhalis  
US-09-540-236-2077

Query Match 69.8%; Score 30; DB 4; Length 231;  
Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQIMDRV 9  
||:||||:  
DB 149 NLQSIIDRI 157

RESULT 17  
US-09-134-000C-6666  
Sequence 6666, Application US/09134000C  
Patent No. 6617156  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
FILE REFERENCE: 032796-032  
CURRENT APPLICATION NUMBER: US/09/134,000C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/055,778  
PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 6666  
LENGTH: 412  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-6666

Query Match 69.8%; Score 30; DB 4; Length 412;  
Best Local Similarity 62.5%; Pred. No. 2.3e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQIMDRV 9  
||:||||:  
DB 375 IQVIMDRV 382

RESULT 18  
US-09-252-991A-29575  
Sequence 29575, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 29575  
LENGTH: 656  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29575

Query Match 69.8%; Score 30; DB 4; Length 656;  
Best Local Similarity 55.6%; Pred. No. 3.8e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQIMDRV 9  
||:||||:  
DB 444 NLNVIIDRI 452

RESULT 19  
US-09-134-001C-4480

; Sequence 4480, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: CTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4480  
; LENGTH: 673  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4480

Query Match 69.8%; Score 30; DB 3; Length 673;  
Best Local Similarity 66.7%; Pred. No. 3.9e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 NLQOLMDRV 9  
DB 473 NQOOLMDRI 481

RESULT 20  
US-09-489-039A-11547  
; Sequence 11547, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 11547  
; LENGTH: 808  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-11547

Query Match 69.8%; Score 30; DB 4; Length 808;  
Best Local Similarity 55.6%; Pred. No. 4.8e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 NLQOLMDRV 9  
DB 409 NFKLMDQI 417

RESULT 21  
US-09-583-110-5094  
; Sequence 5094, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 5094  
; LENGTH: 1116  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-583-110-5094

Query Match 69.8%; Score 30; DB 4; Length 1116;  
Best Local Similarity 85.7%; Pred. No. 6.8e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LQOLMDR 8  
DB 320 LQOLMDR 326

RESULT 22  
US-09-040-738-2  
; Sequence 2, Application US/09040738  
; Patent No. 6207374  
; GENERAL INFORMATION:  
; APPLICANT: Sampson et al.  
; TITLE OF INVENTION: Tubercous Sclerosis 2 Gene and Uses  
; TITLE OF INVENTION: Thereof  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Kathleen M. Williams, Banner & Witcoff,  
; ADDRESSER: Ltd.  
; STREET: One Financial Center  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02111

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Wordperfect 6.1a

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/040,738  
; FILING DATE: Concurrently herewith  
; CLASSIFICATION:

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB9326470.3  
; FILING DATE: 24-December-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB9411900.5  
; FILING DATE: 14-June-1994

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB94/02823  
; FILING DATE: 23-December-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/652,426

FILING DATE: 30-May-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams, Kathleen M.  
; REGISTRATION NUMBER: 34,380  
; REFERENCE/DOCKET NUMBER: 3265/73963  
; TELEPHONE: 617-345-9110  
; TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 2;  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1784 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-040-738-2

Query Match 69.8%; Score 30; DB 3; Length 1784;  
Best Local Similarity 75.0%; Pred. No. 1.1e+03;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 8  
|||  
DB 444 NLOQLMER 451

## RESULT 23

US-08-652-426A-2  
; Sequence 2, Application US/08652426A  
; Patent No. 6232452

GENERAL INFORMATION:

APPLICANT: Sampson et al.

TITLE OF INVENTION: Tubercous Sclerosis 2 Gene and Uses

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kathleen M. Williams, Banner & Witcoff,

STREET: One Financial Center

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02111

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 6.1a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/652,426A

FILING DATE: 01-October-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB9326470.3

FILING DATE: 12/24/93

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB9411900.5

FILING DATE: 06/14/94

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB94/02823

FILING DATE: 12/23/94

ATTORNEY/AGENT INFORMATION:

NAME: Williams, Kathleen M.

REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: 96,738

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-345-9100

TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1784 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-652-426A-2

QY 1 NLOQLMDR 8  
|||  
DB 444 NLOQLMER 451

Query Match 69.8%; Score 30; DB 3; Length 1784;  
Best Local Similarity 75.0%; Pred. No. 1.1e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

## RESULT 24

US-09-270-767-38268

; Sequence 38268, Application US/09270767

; Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 38268

LENGTH: 133

TYPE: PRT

ORGANISM: Drosophila melanogaster

FEATURE:

OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-38268

QY 1 NLOQLMDR 8  
|||  
DB 13 NVRLILDR 20

Query Match 67.4%; Score 29; DB 4; Length 133;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

## RESULT 25

US-09-270-767-53485

; Sequence 53485, Application US/09270767

; Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 53485

LENGTH: 133

TYPE: PRT

ORGANISM: Drosophila melanogaster

FEATURE:

OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-53485

QY 1 NLOQLMDR 8  
|||  
DB 13 NVRLILDR 20

Query Match 67.4%; Score 29; DB 4; Length 133;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

## RESULT 26

US-09-134-001C-4694

; Sequence 4694, Application US/09134001C

; Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 4694

LENGTH: 228

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4694

QY 1 NLOQLMDR 8  
|||  
DB 13 NVRLILDR 20

Query Match 67.4%; Score 29; DB 4; Length 133;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

## RESULT 27

US-09-134-001C-4694

; Sequence 4694, Application US/09134001C

; Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

Query Match 67.4%; Score 29; DB 3; Length 228;  
Best Local Similarity 62.5%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 8  
|:|:|:|:  
40 NLOQLPDK 47

RESULT 27  
US-09-134-001C-5495

; Sequence 5495, Application US/09134001C  
; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 5495

; LENGTH: 228

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5495

Query Match 67.4%; Score 29; DB 3; Length 228;  
Best Local Similarity 62.5%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 8  
|:|:|:|:  
40 NLOQLPDK 47

RESULT 28  
US-09-818-780-93

; Sequence 93, Application US/09818780

; Patent No. 6677146

; GENERAL INFORMATION:

; APPLICANT: McHenry, Charles

; TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III HOLOENZYME

; FILE REFERENCE: 1794.0030004

; CURRENT APPLICATION NUMBER: US/09/818,780

; PRIOR FILING DATE: 2001-03-28

; PRIOR APPLICATION NUMBER: US 60/192,736

; PRIOR FILING DATE: 2000-03-28

; NUMBER OF SEQ ID NOS: 98

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 93

; LENGTH: 269

; TYPE: PRT

; ORGANISM: Rickettsia prowazekii

US-09-818-780-93

Query Match 67.4%; Score 29; DB 4; Length 269;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 9  
|:|:|:|:  
197 NLOQLMDR 205

RESULT 29  
US-09-252-991A-30349  
; Sequence 30349, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 30349

; LENGTH: 280

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30349

Query Match 67.4%; Score 29; DB 4; Length 280;  
Best Local Similarity 71.4%; Pred. No. 2.4e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOQLMD 7  
|:|:|:|:  
72 NLOQLMD 78

RESULT 30  
US-10-067-443-6  
; Sequence 6, Application US/10067443

; Patent No. 6642041

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL METALLOPROTEASE HIGHLY EXPRESSED I

; TITLE OF INVENTION: SPIRAL CORN, MP-1

; FILE REFERENCE: D0073 NP

; CURRENT APPLICATION NUMBER: US/10/067,443

; PRIOR FILING DATE: 2002-02-05

; PRIOR APPLICATION NUMBER: US 60/266,518

; PRIOR FILING DATE: 2001-02-05

; PRIOR APPLICATION NUMBER: US 60/282,814

; PRIOR FILING DATE: 2001-04-10

; NUMBER OF SEQ ID NOS: 71

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Helicobacter pylori

US-10-067-443-6

Query Match 67.4%; Score 29; DB 4; Length 340;  
Best Local Similarity 55.6%; Pred. No. 3e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 9  
|:|:|:|:  
54 NLOQLMDR 62

RESULT 31  
US-09-252-991A-33698  
; Sequence 23698, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23698  
; LENGTH: 369  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23698

Query Match 67.4%; Score 29; DB 4; Length 369;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QLMDR 8  
Db 258 QLMDR 263

RESULT 32  
US-09-252-991A-30109  
; Sequence 30109, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30109  
; LENGTH: 420  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30109

Query Match 67.4%; Score 29; DB 4; Length 420;  
Best Local Similarity 66.7%; Pred. No. 3.7e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLQMDRV 9  
Db 290 NLQMDRV 298

RESULT 33  
US-09-248-796A-15391  
; Sequence 15391, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; PRIOR FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 15391  
; LENGTH: 508  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-15391

Query Match 67.4%; Score 29; DB 4; Length 508;  
Best Local Similarity 75.0%; Pred. No. 4.6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LQMDRV 9  
Db 6 LQMDRL 13

RESULT 34  
US-09-252-991A-22687  
; Sequence 22687, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 22687  
; LENGTH: 509  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22687

Query Match 67.4%; Score 29; DB 4; Length 509;  
Best Local Similarity 85.7%; Pred. No. 4.6e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQMDRV 8  
Db 116 LQMDRV 122

RESULT 35  
US-09-252-991A-18744  
; Sequence 18744, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18744  
; LENGTH: 573  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18744

Query Match 67.4%; Score 29; DB 4; Length 573;  
Best Local Similarity 66.7%; Pred. No. 5.3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NLQMDRV 9  
Db 122 NLQMDRV 130

RESULT 36  
US-09-252-991A-29441  
; Sequence 29441, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 29441  
LENGTH: 575  
TYPE: prt  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29441

Query Match 67.4%; Score 29; DB 4; Length 575;  
Best Local Similarity 66.7%; Pred. No. 5.3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QILMDR 9  
DB 191 QILMDR 199

RESULT 37  
US-08-871-266B-18  
Sequence 18, Application US/08871266B  
Patent No. 5871991  
GENERAL INFORMATION:  
APPLICANT: Elrod, Susan L.  
TITLE OF INVENTION: Aspergillus Oryzae 5-Aminolevulinic  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSER: No. 5871991disk Of No. 5871991th America, Inc.  
STREET: 405 Lexington Avenue - 64th Fl.  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/871,266B  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-878-9652  
TELEFAX: 212-878-9655  
TELEX:  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 587 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 5871991e  
US-08-871-266B-18

Query Match 67.4%; Score 29; DB 2; Length 587;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QILMDR 8  
DB 459 QILMDR 464

RESULT 38  
US-09-018-864A-18  
Sequence 18, Application US/0901864A  
Patent No. 5958747  
GENERAL INFORMATION:  
APPLICANT: Elrod, Susan L.  
TITLE OF INVENTION: Aspergillus Oryzae 5-Aminolevulinic  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSER: No. 5958747disk Of No. 5958747th America, Inc.  
STREET: 405 Lexington Avenue - 64th Fl.  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/018,864A  
FILING DATE:  
CLASSIFICATION: 08/871,266  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/871,266  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lambiris, Elias J  
REGISTRATION NUMBER: 33,728  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-878-9652  
TELEFAX: 212-878-9655  
TELEX:  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 587 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 5958747e  
US-09-018-864A-18

Query Match 67.4%; Score 29; DB 2; Length 587;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QILMDR 8  
DB 459 QILMDR 464

RESULT 39  
US-08-871-267B-24  
Sequence 24, Application US/08871267B  
Patent No. 6100057  
GENERAL INFORMATION:  
APPLICANT: Elrod, Susan L.  
APPLICANT: Cherry, Joel R.  
APPLICANT: Jones, Aubrey  
TITLE OF INVENTION: A Method for Increasing Hemoprotein  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESSER: No. 6100057disk Of No. 6100057th America, Inc.

STREET: 405 Lexington Avenue - 64th Fl.  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/871,267B  
FILING DATE: 9-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rozek, Carol E.  
REGISTRATION NUMBER: 36,993  
REFERENCE/DOCKET NUMBER: 4771.200-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-878-9652  
TELEFAX: 212-878-9655  
TELEX:  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 587 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6100057e  
US-08-871-267B-24

Query Match 67.4%; Score 29; DB 3; Length 587;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 QLMDR 8  
DB 459 QLMDR 464

RESULT 40  
US-09-618-419-24  
Sequence 24, Application US/09618419  
Patent No. 6261827  
GENERAL INFORMATION:  
APPLICANT: Elrod, Susan L.  
Cherry, Joel R.  
Jones, Aubrey  
TITLE OF INVENTION: A Method for Increasing Hemoprotein  
Production in Filamentous Fungi  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESS: No. 62618270 No. 6261827disk Of No. 6261827th America, Inc.  
STREET: 405 Lexington Avenue - 64th Fl.  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10174  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/618,419  
FILING DATE: 18-Jul-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/871,267  
FILING DATE: 9-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Rozek, Carol E.  
REGISTRATION NUMBER: 36,993  
REFERENCE/DOCKET NUMBER: 4771.200-US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-878-9652  
TELEFAX: 212-878-9655  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 587 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: No. 6261827e  
SEQUENCE DESCRIPTION: SEQ ID NO: 24:  
US-09-618-419-24

Query Match 67.4%; Score 29; DB 3; Length 587;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 QLMDR 8  
DB 459 QLMDR 464

Search completed: January 12, 2005, 20:17:42  
Job time : 22.6 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2005, 20:16:00 ; Search time 93.6 Seconds  
(without alignments)  
34.739 Million cell updates/sec

Title: US-09-870-216C-11

Perfect score: 43

Sequence: 1 NQLQMLMDRV 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*\n2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*\n3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*\n4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*\n5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*\n6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*\n7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*\n8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*\n9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*\n10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*\n11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*\n12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*\n13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*\n14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*\n15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*\n16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*\n17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*\n18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*\n19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*\n20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	9	11	US-09-870-216C-11
2	43	100.0	9	13	US-10-017-327-11
3	43	100.0	256	9	US-09-925-300-1353
4	43	100.0	352	11	US-09-870-216C-2
5	43	100.0	352	13	US-10-017-327-2
6	43	100.0	352	16	US-10-408-765A-347
7	36	83.7	311	17	US-10-739-930-6423
8	35	81.4	180	15	US-10-264-237-1883
9	35	81.4	239	9	US-09-925-300-1329
10	34	79.1	371	17	US-10-425-115-239256
11	34	79.1	372	15	US-10-425-114-59828
12	33	76.7	874	14	US-10-163-214-13
13	33	76.7	987	16	US-10-437-963-113606

14	32	74.4	125	9	US-09-867-550-422	Sequence 422, App
15	32	74.4	155	17	US-10-425-115-357844	Sequence 357844,
16	32	74.4	178	15	US-10-424-599-233102	Sequence 233102,
17	32	74.4	203	14	US-10-369-493-11994	Sequence 11994, A
18	32	74.4	423	16	US-10-437-963-166526	Sequence 166526,
19	32	74.4	2697	9	US-09-961-527A-5	Sequence 5, Appli
20	32	74.4	4080	15	US-10-307-817-138	Sequence 138, App
21	32	74.4	4624	16	US-10-408-765A-2991	Sequence 2991, Ap
22	31	72.1	156	16	US-10-437-963-113390	Sequence 113390,
23	31	72.1	160	17	US-10-425-115-325709	Sequence 325709,
24	31	72.1	175	16	US-10-437-963-142657	Sequence 142657,
25	31	72.1	175	16	US-10-437-963-125323	Sequence 125323,
26	31	72.1	229	17	US-10-425-115-218767	Sequence 218767,
27	31	72.1	270	9	US-09-738-626-5133	Sequence 5133, Ap
28	31	72.1	270	15	US-10-627-476-676	Sequence 676, App
29	31	72.1	320	15	US-10-425-114-63198	Sequence 63198, A
30	31	72.1	336	17	US-10-425-115-345058	Sequence 345058,
31	31	72.1	455	17	US-10-739-930-10942	Sequence 10942, A
32	31	72.1	529	17	US-10-848-111-4	Sequence 11, Appli
33	31	72.1	542	15	US-10-289-762-496	Sequence 496, App
34	31	72.1	823	14	US-10-425-115-340748	Sequence 340748,
35	31	72.1	833	14	US-10-369-493-14960	Sequence 14960, A
36	31	72.1	838	14	US-10-369-493-11738	Sequence 11738, A
37	31	72.1	838	14	US-10-369-493-14794	Sequence 14794, A
38	31	72.1	916	15	US-10-287-426-330	Sequence 330, App
39	31	72.1	1504	14	US-10-369-493-22466	Sequence 22466, A
40	30	69.8	55	15	US-10-335-977-6139	Sequence 6139, Ap
41	30	69.8	130	17	US-10-425-115-362529	Sequence 362529,
42	30	69.8	132	14	US-10-104-047-2471	Sequence 2471, Ap
43	30	69.8	161	16	US-10-767-701-50312	Sequence 50312, A
44	30	69.8	167	14	US-10-183-687-390	Sequence 390, App
45	30	69.8	181	17	US-10-425-115-331931	Sequence 331931,

## ALIGNMENTS

RESULT 1  
US-09-870-216C-11  
Sequence 11, Application US/09870216C  
Publication NO. US20040138135A1  
GENERAL INFORMATION:  
APPLICANT: Charles A. Nicolette  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
FILE REFERENCE: 68126881210100  
CURRENT APPLICATION NUMBER: US/09/870,216C  
CURRENT FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 60/209,391  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: 60/226,256  
PRIOR FILING DATE: 2000-08-17  
PRIOR APPLICATION NUMBER: 60/257,008  
PRIOR FILING DATE: 2000-12-20  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-870-216C-11  
Query Match 100.0%; Score 43; DB 11; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0;  
QY 1 NQLQMLMDRV 9  
DB 1 NQLQMLMDRV 9  
RESULT 2  
US-10-017-327-11  
Sequence 11, Application US/10017327

```
Publication No. US20020155471A1
GENERAL INFORMATION:
APPLICANT: Charles A. Nicolette
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND
METHODS FOR USING SAME
FILE REFERENCE: G2 2101.20
CURRENT APPLICATION NUMBER: US/10/017,327
CURRENT FILING DATE: 2001-12-06
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
US-10-017-327-11

Query Match      100.0%; Score 43; DB 13; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NLQQLMDRV 9
DB      1 NLQQLMDRV 9

RESULT 3
US-09-925-300-1353
Sequence 1353, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben,
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1353
LENGTH: 256
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1353

Query Match      100.0%; Score 43; DB 9; Length 256;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NLQQLMDRV 9
DB      146 NLQQLMDRV 154

RESULT 4
US-09-870-216C-2
Sequence 2, Application US/09870216C
Publication No. US20040138135A1
GENERAL INFORMATION:
APPLICANT: Charles A. Nicolette
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER
FILE REFERENCE: 68126881210100
CURRENT APPLICATION NUMBER: US/09/870,216C
CURRENT FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 60/209,391
PRIOR FILING DATE: 2000-05-31
```

```
PRIOR APPLICATION NUMBER: 60/226,256
PRIOR FILING DATE: 2000-08-17
PRIOR APPLICATION NUMBER: 60/257,008
PRIOR FILING DATE: 2000-12-20
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 352
TYPE: PRT
ORGANISM: Homo sapiens
US-09-870-216C-2

Query Match      100.0%; Score 43; DB 11; Length 352;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NLQQLMDRV 9
DB      242 NLQQLMDRV 250

RESULT 5
US-10-017-327-2
Sequence 2, Application US/10017327
Publication No. US20020155471A1
GENERAL INFORMATION:
APPLICANT: Charles A. Nicolette
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND
METHODS FOR USING SAME
FILE REFERENCE: G2 2101.20
CURRENT APPLICATION NUMBER: US/10/017,327
CURRENT FILING DATE: 2001-12-06
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 352
TYPE: PRT
ORGANISM: Homo sapiens
US-10-017-327-2

Query Match      100.0%; Score 43; DB 13; Length 352;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NLQQLMDRV 9
DB      242 NLQQLMDRV 250

RESULT 6
US-10-408-765A-347
Sequence 347, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fany, Eoin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Warnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
FILE REFERENCE: 660088.465
CURRENT APPLICATION NUMBER: US/10/408,765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 347
LENGTH: 352
TYPE: PRT
ORGANISM: Homo sapiens
US-10-408-765A-347
```

Query Match 100.0%; Score 43; DB 16; Length 352;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLOQLMDRV 9  
Db 242 NLOQLMDRV 250

RESULT 7  
US-10-739-930-6423  
; Sequence 6423, Application US/10739930  
; Publication No. US20040216190A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; FILE REFERENCE: 38-21(53377)B  
; CURRENT APPLICATION NUMBER: US/10/739,930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO 6423  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C3789\_1.p  
US-10-739-930-6423

Query Match 83.7%; Score 36; DB 17; Length 311;  
Best Local Similarity 87.5%; Pred. No. 62;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQQLMDRV 9  
Db 146 LQQLMDRV 153

RESULT 8  
US-10-264-237-1883  
; Sequence 1883, Application US/10264237  
; Publication No. US2004009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA131P1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 1883  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-237-1883

Query Match 81.4%; Score 35; DB 15; Length 180;  
Best Local Similarity 87.5%; Pred. No. 55;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLOQLMDRV 8  
Db 23 NLOQLMDRV 30

RESULT 9  
US-09-925-300-1229  
; Sequence 1229, Application US/09925300

; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1229  
; LENGTH: 239  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-300-1229

Query Match 81.4%; Score 35; DB 9; Length 239;  
Best Local Similarity 87.5%; Pred. No. 74;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLOQLMDRV 8  
Db 82 NLOQLMDRV 89

RESULT 10  
US-10-425-115-239256  
; Sequence 239256, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 239256  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_14978C.1.pcp  
US-10-425-115-239256

Query Match 79.1%; Score 34; DB 17; Length 371;  
Best Local Similarity 77.8%; Pred. No. 1.9e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLOQLMDRV 9  
Db 356 NLOQLMDRV 364

RESULT 11  
US-10-425-114-59828  
; Sequence 59828, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59828
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3911-002-C10_F11.pdp
US-10-425-114-59828

Query Match          79.1%; Score 34; DB 15; Length 372;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 NLQILMDRV 9
Db      357 NAKLMDRV 365

RESULT 12
US-10-163-214-13
; Sequence 13, Application US/10163214
; Publication No. US20030097688A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Brogile, Karen E.
; APPLICANT: Butler, Karlene H.
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Starch Synthase Isoform V
; FILE REFERENCE: BR1520 US NA
; CURRENT APPLICATION NUMBER: US/10/163,214
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/297,099
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Vigna unguiculata
US-10-163-214-13

Query Match          76.7%; Score 33; DB 14; Length 874;
Best Local Similarity 75.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 NLQILMDR 8
Db      195 NLQILMDK 202

RESULT 13
US-10-437-963-113606
; Sequence 113606, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
```

```

; SEQ ID NO 113606
; LENGTH: 987
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_17378C.1.pdp
US-10-437-963-113606

Query Match          76.7%; Score 33; DB 16; Length 987;
Best Local Similarity 66.7%; Pred. No. 8.3e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 NLQILMDRV 9
Db      564 DLQILVDRI 572

RESULT 14
US-09-867-550-422
; Sequence 422, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Menrabad, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 422
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-422

Query Match          74.4%; Score 32; DB 9; Length 125;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 NLQILMDR 8
Db      22 NLQILMDQ 29

RESULT 15
US-10-425-115-357844
; Sequence 357844, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 357844
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(155)
```

OTHER INFORMATION: unsure at all Xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: MRT4577\_89523C.1.pep  
US-10-425-115-357844

Query Match 74.4%; Score 32; DB 17; Length 155;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NLQILMDRV 9  
Db 144 NLILFMDRI 152

RESULT 16  
US-10-424-599-233102  
Sequence 233102, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 233102  
LENGTH: 178  
TYPE: PRT  
ORGANISM: Glycine max

FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_52515C.1.pep  
US-10-424-599-233102

Query Match 74.4%; Score 32; DB 15; Length 178;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLQILMDR 8  
Db 76 NLQILVMDQ 83

RESULT 17  
US-10-369-493-11994

Sequence 11994, Application US/10369493  
Publication No. US2003023675A1  
GENERAL INFORMATION:

APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 11994  
LENGTH: 203  
TYPE: PRT  
ORGANISM: Mesorhizobium loti  
US-10-369-493-11994

Query Match 74.4%; Score 32; DB 14; Length 203;  
Best Local Similarity 87.5%; Pred. No. 2.4e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LQILMDRV 9  
Db 125 LDILMDRV 132

RESULT 18  
US-10-437-963-166526  
Sequence 166526, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 166526  
LENGTH: 423  
TYPE: PRT  
ORGANISM: Oryza sativa

FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_65226C.1.pep  
US-10-437-963-166526

Query Match 74.4%; Score 32; DB 16; Length 423;  
Best Local Similarity 66.7%; Pred. No. 5.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLQILMDRV 9  
Db 354 NLQILMDRL 362

RESULT 19  
US-09-961-527A-5

Sequence 5, Application US/09961527A  
Patent No. US20020142324A1  
GENERAL INFORMATION:

APPLICANT: Wang, Xun  
APPLICANT: Turgeson, B. Gillian  
APPLICANT: Yoder, Olen  
APPLICANT: Mu, Jiansuo  
TITLE OF INVENTION: Fungal target genes and methods to identify those genes  
FILE REFERENCE: TW0129-UT  
CURRENT APPLICATION NUMBER: US/09/961,527A  
CURRENT FILING DATE: 2001-09-24  
PRIOR APPLICATION NUMBER: US 60/234,673  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: US 60/234,650  
PRIOR FILING DATE: 2000-09-22  
NUMBER OF SEQ ID NOS: 19  
SEQ ID NO 5  
LENGTH: 2697  
TYPE: PRT  
ORGANISM: Cochliobolus  
US-09-961-527A-5

Query Match 74.4%; Score 32; DB 9; Length 2697;  
Best Local Similarity 75.0%; Pred. No. 3.8e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LQILMDRV 9  
Db 605 LQILMDRI 612

```
RESULT 20
US-10-307-817-138
; Sequence 138, Application US/10307817
; Publication No. US20040058338A1
; GENERAL INFORMATION:
; APPLICANT: Agee et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-502C
; CURRENT APPLICATION NUMBER: US/10/307,817
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 138
; LENGTH: 4080
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-307-817-138

Query Match
Best Local Similarity 74.4%; Score 32; DB 15; Length 4080;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQQLMDRV 9
DB 796 DLELLMDRV 804

RESULT 21
US-10-408-765A-2991
; Sequence 2991, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary W.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 66088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2991
; LENGTH: 4624
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2991

Query Match
Best Local Similarity 74.4%; Score 32; DB 16; Length 4624;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQQLMDRV 9
DB 812 DLELLMDRV 820

RESULT 22
US-10-437-963-113390
; Sequence 113390, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
US-10-437-963-113390

Query Match
Best Local Similarity 72.1%; Score 31; DB 16; Length 156;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQQLMDRV 9
DB 94 NVELSMDRV 102

RESULT 23
US-10-425-115-325709
; Sequence 325709, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 325709
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: MRT4577_60115C.1.pep
US-10-425-115-325709

Query Match
Best Local Similarity 72.1%; Score 31; DB 17; Length 160;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQQLMDRV 9
DB 151 LQQLMDRV 158

RESULT 24
US-10-437-963-142657
; Sequence 142657, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

```
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 142657
LENGTH: 162
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(162)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_43642C.1.pep
US-10-437-963-142657

Query Match
Best Local Similarity 72.1%; Score 31; DB 16; Length 162;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9
Db 20 NVELSMDRV 28

RESULT 25
US-10-437-963-125323
Sequence 125323, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Mu, Wei
APPLICANT: Bouharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 125323
LENGTH: 175
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_27979C.1.pep
US-10-437-963-125323

Query Match
Best Local Similarity 72.1%; Score 31; DB 16; Length 175;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9
Db 80 NLPPLMDRV 88

RESULT 26
US-10-425-115-218767
Sequence 218767, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Mu, Wei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
```

```
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 218767
LENGTH: 229
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(229)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_131106C.1.pep
US-10-425-115-218767

Query Match
Best Local Similarity 72.1%; Score 31; DB 17; Length 229;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9
Db 138 NLSLMDRV 146

RESULT 27
US-09-738-626-5133
Sequence 5133, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5133
LENGTH: 270
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5133

Query Match
Best Local Similarity 72.1%; Score 31; DB 9; Length 270;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9
Db 22 NLPPLMDRV 30

RESULT 28
US-10-627-476-676
Sequence 676, Application US/10627476
Publication No. US20040030116A1
GENERAL INFORMATION:
APPLICANT: Pompejus, Mark
APPLICANT: Krieger, Burkhard
```

```

APPLICANT: Schoder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauser, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
TITLE OF INVENTION: TRANSPORT
FILE REFERENCE: BG1-125CPN
CURRENT APPLICATION NUMBER: US/10/627,476
CURRENT FILING DATE: 2003-07-25
PRIOR APPLICATION NUMBER: 09/602,787
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US5N 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: DE 19931454.3
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931478.0
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931563.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932122.1
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932124.8
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932128.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932180.9
PRIOR FILING DATE: 1999-07-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 678
SEQ ID NO 676
LENGTH: 270
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-10-627-476-676

Query Match      72.1%;   Score 31;   DB 15;   Length 270;
Best Local Similarity 66.7%;   Pred. No. 5.2e+02;
Matches 6;   Conservative 2;   Mismatches 1;   Indels 0;   Gaps 0;

CY      1  NQLGLMDRV 9
Db      22  NLEPLIDRV 30

RESULT 29
US-10-425-114-63198
; Sequence 63198, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63198
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17286D01_F11.pep
US-10-425-114-63198

Query Match      72.1%;   Score 31;   DB 15;   Length 320;
Best Local Similarity 77.8%;   Pred. No. 6.2e+02;

```

```

Matches      7;  Conservative      0;  Mismatches      2;  Indels      0;  Gaps      0;

QY      1  NLQOLMDRV  9
      || || || ||
Db      122  NLPLLYDRV  130

RESULT 30
US-10-425-115-345058
; Sequence 345058, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Roca, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 345058
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_7785C.1.pep
US-10-425-115-345058

Query Match      72.1%;  Score 31;  DB 17;  Length 336;
Best Local Similarity 77.8%;  Pred. No. 6.6e+02;
Matches      7;  Conservative      0;  Mismatches      2;  Indels      0;  Gaps      0;

QY      1  NLQOLMDRV  9
      || || || ||
Db      138  NLPLLYDRV  146

RESULT 31
US-10-739-930-10942
; Sequence 10942, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 10942
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C8746_1.P
US-10-739-930-10942

Query Match      72.1%;  Score 31;  DB 17;  Length 455;
Best Local Similarity 77.8%;  Pred. No. 9e+02;
Matches      7;  Conservative      0;  Mismatches      2;  Indels      0;  Gaps      0;

QY      1  NLQOLMDRV  9
      || || || ||
Db      266  NLSTLMDRV  274

RESULT 32
US-10-848-111-4
; Sequence 4, Application US/10848111
; Publication No. US20040235107A1
; GENERAL INFORMATION:

```



```

; APPLICANT: Rosenberg, Eugene
; APPLICANT: Ron, Eljora
; APPLICANT: Orr, Elisha
; APPLICANT: Paltan, Yoasi
; TITLE OF INVENTION: GENE CLUSTER
; FILE REFERENCE: 27757
; CURRENT APPLICATION NUMBER: US/10/848,111
; CURRENT FILING DATE: 2004-05-19
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 4
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-10-848-111-4

Query Match
Best Local Similarity 72.1%; Score 31; DB 17; Length 529;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LQIUMDRV 9
Db 48 LQIUMDRV 55

RESULT 33
US-10-289-762-496
; Sequence 496, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffla, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 496
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-10-289-762-496

Query Match
Best Local Similarity 72.1%; Score 31; DB 15; Length 542;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLQIUMDRV 9
Db 531 NLQIUMDRV 539

RESULT 34
US-10-425-115-340748
; Sequence 340748, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 340748
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
```

```

; OTHER INFORMATION: Clone ID: MRT4577_73932C.1.pep
; US-10-425-115-340748

Query Match
Best Local Similarity 72.1%; Score 31; DB 17; Length 823;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQIUMDRV 9
Db 573 NLQIUMDRV 581

RESULT 35
US-10-369-493-14960
; Sequence 14960, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14960
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
; US-10-369-493-14960

Query Match
Best Local Similarity 72.1%; Score 31; DB 14; Length 833;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQIUMDRV 9
Db 426 LQIUMDRV 433

RESULT 36
US-10-369-493-11738
; Sequence 11738, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11738
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
; US-10-369-493-11738

Query Match
Best Local Similarity 72.1%; Score 31; DB 14; Length 838;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      2 LQIIMDRV 9
Db      430 IQLLVDRV 437

RESULT 37
US-10-369-493-14794
; Sequence 14794, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 14794
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-14794

Query Match      72.1%; Score 31; DB 14; Length 838;
Best Local Similarity 75.0%; Pred. No. 1.7e+03;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 LQIIMDRV 9
Db      430 IQLLVDRV 437

RESULT 38
US-10-287-226-330
; Sequence 330, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.
; APPLICANT: Alsobrook, John P.
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Ference,
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Chant, John S.
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eisen, Andrew,
; APPLICANT: Elleman, Karen,
; APPLICANT: Gangolli, Esba A.
; APPLICANT: Gorman, Linda
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khrantsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malvankar, Utiel M.
; APPLICANT: Macdougall, John R.
; APPLICANT: Mezes, Peter S.
; APPLICANT: Miller, Charles E.
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Paturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Rothenberg, Mark B.

```

```

; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Spytek, Kimberley A.
; APPLICANT: Taupier, Jr., Raymond J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zehnisen, Bryan D.
; APPLICANT: Zhong, Wei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,409
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/364,227
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/334,027
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: Cursesqlist version 0.1
; SEQ ID NO 330
; LENGTH: 916
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-226-330

Query Match      72.1%; Score 31; DB 15; Length 916;
Best Local Similarity 75.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 NLQIIMDR 8
Db      740 HLQIIMDK 747

RESULT 39
US-10-369-493-22466
; Sequence 22466, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22466
; LENGTH: 1504
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-22466

```

Query Match 72.1%; Score 31; DB 14; Length 1504;  
 Best Local Similarity 75.0%; Pred. No. 3.2e+03;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOLLMDR 8  
 :||||:|  
 Db 237 DLQLLDR 244

RESULT 40  
 US-10-335-977-6139  
 ; Sequence 6139, Application US/10335977  
 ; Publication No. US20040052799A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DOUGLAS SMITH et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
 ; RELATING TO HELICOBACTER PYLORI FOR  
 ; DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 10031  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSER: LAHYE & COCKFIELD  
 ; STREET: 28 State Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109-1875  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD-ROM ISO9660  
 ; COMPUTER: IBM PC Compatible  
 ; OPERATING SYSTEM: Windows NT 4.0  
 ; SOFTWARE: UNIX  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/335,977  
 ; FILING DATE: 30-Dec-2002  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/993,002  
 ; FILING DATE: 17-DEC-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Mandragouras, Amy E.  
 ; REGISTRATION NUMBER: 36,207  
 ; REFERENCE/DOCKET NUMBER: GTN-018  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617)227-7400  
 ; TELEFAX: (617)742-4214  
 ; INFORMATION FOR SEQ ID NO: 6139:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 55 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHEICAL: YES  
 ; ORGANISM: Helicobacter pylori  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (B) LOCATION 1...55  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6139:  
 ; US-10-335-977-6139

Query Match 69.8%; Score 30; DB 15; Length 55;  
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLOLLMD 7  
 :||||:|  
 Db 45 NLOLLMD 51

Search completed: January 12, 2005, 20:48:30  
 Job time : 96 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:37:43 ; Search time 201.4 Seconds

(without alignments)  
16.031 Million cell updates/sec

Title: US-09-870-216c-11

Sequence: 1 NLQILMDRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 35872929 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: geneeqp1980s:\*
- 2: geneeqp1990s:\*
- 3: geneeqp2000s:\*
- 4: geneeqp2001s:\*
- 5: geneeqp2002s:\*
- 6: geneeqp2003as:\*
- 7: geneeqp2003bs:\*
- 8: geneeqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	9	5 ABB08365	Abb08365 The natur
2	43	100.0	9	7 ABR82217	Abt82217 Human ant
3	43	100.0	256	3 AAB56775	Aab56775 Human pro
4	43	100.0	302	7 ADB29060	Adb29060 Human eIF
5	43	100.0	347	7 ADR29059	Adr29059 Human eIF
6	43	100.0	352	5 ABR08360	Abb08360 Human eIF
7	43	100.0	352	7 ABR82212	Abt82212 Human ant
8	43	100.0	352	7 ADJ68541	Adj68541 Human hea
9	43	100.0	368	7 ADC31312	Adc31312 Human nov
10	36	83.7	260	3 AAG57646	Aag57646 Arabidops
11	36	83.7	311	3 AAG57645	Aag57645 Arabidops
12	36	83.7	334	3 AAB57644	Aab57644 Arabidops
13	35	81.4	133	4 AAB92665	Aab92665 Human pro
14	35	81.4	152	3 AAG01149	Aag01149 Human sec
15	35	81.4	180	2 AAY28922	Aay28922 Human reg
16	35	81.4	180	5 ABB89507	Abb89507 Human pol
17	35	81.4	180	7 ADD22442	Add22442 HLA-B46 T
18	35	81.4	180	7 ADI15876	Adi15876 Human PP
19	35	81.4	180	7 ADI15933	Adi15933 Human PP
20	35	81.4	180	8 ADJ81699	Adj81699 Tumour an
21	35	81.4	239	3 AAB56651	Aab56651 Human pro
22	35	81.4	1809	7 ADB55015	Adb55015 Rat Prote
23	35	81.4	1809	7 ADB55009	Adb55009 Rat Prote
24	35	81.4	1809	7 ADB55006	Adb55006 Rat Prote
25	35	81.4	1809	7 ADB55012	Adb55012 Rat Prote

26	34	79.1	1116	4 ABB59885	Abb59885 Drosophil
27	33	76.7	509	7 ADM26506	Adm26506 Hyperther
28	33	76.7	874	7 AAE39667	Aae39667 Vigna ung
29	32	74.4	125	5 ABB63841	Abb63841 Human ORF
30	32	74.4	228	4 AAU44957	Aau44957 Propionib
31	32	74.4	228	6 ABM41476	Abm41476 Propionib
32	32	74.4	269	5 ABB28324	Abb28324 Streptoco
33	32	74.4	283	8 ADR46631	Adr46631 Streptoco
34	32	74.4	621	5 ABB28470	Abb28470 Streptoco
35	32	74.4	2697	6 AAO26550	Aao26550 Cocchioba
36	32	74.4	4080	8 ADO42289	Ado42289 Human NOV
37	32	74.4	4624	7 ADE78572	Ade78572 Dynamin ax
38	32	74.4	4624	7 ADJ71185	Adj71185 Human hea
39	31	72.1	152	8 ADR48512	Adr48512 Streptoco
40	31	72.1	270	4 AAB76847	Aab76847 Corynebac
41	31	72.1	41	7 AAG91379	Aag91379 C glutami
42	31	72.1	276	7 ABO69260	Abog69260 Pseudomon
43	31	72.1	309	3 AAY57447	Aay57447 Mouse hom
44	31	72.1	403	8 ADM80057	Adm80057 Spliramyct
45	31	72.1	403	8 ADN97573	Adn97573 S ambotac

## ALIGNMENTS

RESULT 1	AB08365	standard; protein; 9 AA.
ID	AB08365	standard; protein; 9 AA.
XX	AB08365;	
AC	AB08365;	
DT	07-MAY-2002	(first entry)
XX		
DE	The natural epitope of human cancer antigen eIF3.	
XX		
KM	Human; melanoma antigen eukaryotic initiation factor 3; eIF3;	
KW	ovarian cancer; MHC; cytosolic; immunomodulator; immune effector cell;	
XX	anti-cancer; vaccine.	
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	Domain	1 /note= "HLA-2 binding residue"
FT	Domain	2 /note= "HLA-2 binding residue"
FT	Domain	3..8 /note= "T-cell receptor (TCR) binding domain"
FT	Domain	9 /note= "HLA-2 binding residue"
XX		
PD	MO200192307-A2.	
XX	06-DEC-2001.	
PF	30-MAY-2001; 2001WO-US017456.	
XX		
PR	31-MAY-2000; 2000US-0209391P.	
PR	17-AUG-2000; 2000US-0226258P.	
PR	20-DEC-2000; 2000US-0257008P.	
XX		
PA	(GENZ ) GENZYME CORP.	
XX		
PI	Nicolette CA;	
XX		
DR	WPI; 2002-139606/18.	
XX	N-PSDB; ABA97216.	
PT	New therapeutic compounds useful against human ovarian cancer, for	
PT	modulating immune response in a subject, and for generating antibodies	
PT	that specifically recognize and bind to these molecules.	
XX		
PS	Claim 10; Page 59; 68pp; English.	

XX The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen B1F3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents the natural epitope of human cancer antigen eIF3 (corresponds  
 CC to residues 242-250 of ABR88360)

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 43; DB 5; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLQILMDRV 9  
 |||||  
 1 NLQILMDRV 9

RESULT 2

ID ABR82217 standard; peptide; 9 AA.

AC ABR82217;

DT 13-OCT-2003 (first entry)

XX Human antigen eIF3 epitope (residues 242-250).

XX Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;

XX cytosol; gene therapy; human; antigen; epitope.

OS Homo sapiens.

XX WO2003050543-A1.

XX 19-JUN-2003 ✓

XX 05-DEC-2001; 2001WO-US047997.

XX 05-DEC-2001; 2001WO-US047997.

XX (GENZ ) GENZYME CORP.

XX Nicolette CA;

XX WPI; 2003-532936/50.

XX Aiding in the diagnosis of a neoplastic condition, useful for treating  
 PT cancer and related malignancies comprises determining the amount of  
 PT expression of an eIF3 protein in a test sample isolated from the cell or  
 PT tissue.

XX Claim 6; Page 30; 77pp; English.

XX The invention relates to aiding in the diagnosis of a neoplastic  
 CC condition or susceptibility to a neoplastic condition of an animal cell  
 CC or tissue. The method involves determining the amount of expression of an  
 CC eukaryotic translation initiation factor 3 (eIF3) protein in a test  
 CC sample isolated from the cell or tissue, and diagnosing a neoplastic  
 CC condition or susceptibility to a neoplastic condition based on the amount

CC of expression of the eIF3 protein. The methods, compounds and kits are  
 CC useful in therapeutics, diagnostic and screening methods for human cancer  
 CC and related malignancies, e.g. ovarian, breast, lung, colon, prostate,  
 CC pancreatic or gastrointestinal cancer, or melanoma. The present sequence  
 CC represents the human antigen eIF3 epitope

XX SQ Sequence 9 AA;

Query Match 100.0%; Score 43; DB 7; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLQILMDRV 9  
 |||||  
 1 NLQILMDRV 9

RESULT 3  
 ID AAB56775 standard; protein; 256 AA.

AC AAB56775;

DT 13-MAR-2001 (first entry)

XX Human prostate cancer antigen protein sequence SEQ ID NO:1353.

XX Human prostate cancer antigen protein sequence SEQ ID NO:1353.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

XX neuroprotective; cytostatic; cardiostatic; immunomodulatory; muscular;

XX vulnereary; gastrointestinal; nephrotoxic; antineoplastic; gynaecological;

XX antibacterial; gene therapy; neutral; immune; reproductive; renal;

XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

XX wound; infectious disease.

XX Homo sapiens.

XX WO200055174-A1.

XX 08-MAR-2000; 2000WO-US005988.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587513/55.

XX N-PSDB; AAF15978.

XX Prostate cancer associated gene sequences, referred to as prostate cancer

XX antigens, useful for treatment, prevention, and diagnosis of disorders

XX such as prostate cancer.

XX Claim 11; Page 1777-1778; 2338pp; English.

XX AAF15566 to AAF15505 encode the human prostate cancer associated

XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

XX The prostate cancer antigens can have neuroprotective, cytostatic,

XX cardiostatic, immunomodulatory, muscular, vulnereary, gastrointestinal,

XX nephrotoxic, antineoplastic, gynaecological and antibacterial activities,

XX and can be used in gene therapy. The prostate cancer antigen

XX polynucleotides may be used for detection of prostate cancer, chromosome

XX identification, as chromosome markers, and for numerous other diagnostic

XX or research purposes. The prostate cancer antigens may be used to treat

XX disorders such as neural, immune, muscular, reproductive,

XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

XX disorders, wounds, and infectious diseases. AAF15506 to AAF16514 to

XX AAB57303 represent sequences used in the exemplification of the present

XX invention

SQ Sequence 256 AA;

Query Match 100.0%; Score 43; DB 3; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9  
 |||||  
 146 NLOQLMDRV 154

RESULT 4

ADBE29060  
 ID ADE29060 standard; protein; 302 AA.

AC ADE29060;

DT 29-JAN-2004 (first entry)

DE Human EIF3S3 selected interacting domain protein - SEQ ID 46.

XX anti-HIV; SID; selected interacting domain; HIV; gene therapy; siRNA;  
 XX silencing RNA; HIV-1 integrase; LEDGF; MCM7; HBO1; Snurportin; VBP1;  
 KW Transportin-SR; EIF3S3; human.

XX Homo sapiens.

XX WO2003046176-A2.

XX PD 05-JUN-2003. X

XX PF 26-NOV-2002; 2002WO-EP013868.

XX PR 26-NOV-2001; 2001US-0333346P.

XX PR 31-MAY-2002; 2002US-0385132P.

XX (HYBR-) HYBRIGENICS.

XX PI Legrain P, Rain J, Benarous R, Emilian S, Berlioz-Torrent C;

XX PI Bloc G;

XX DR WPI: 2003-505199/47.

XX DR N-PSDB; ADE29037.

XX PT New complex between two interacting proteins, useful for screening

XX PT molecules that inhibit human immunodeficiency virus or for preparing a

XX PT medicament for treating HIV-1.

XX PS Claim 9; SEQ ID NO 46; 102pp; English.

XX SQ The invention relates to a novel complex between two interacting proteins

CC listed within the specification. The complex of the invention  
 CC demonstrates anti-HIV activity whilst the SID (selected interacting  
 CC domain) and polypeptides may be useful for screening molecules that  
 CC inhibit human immunodeficiency virus (HIV), as well as during gene  
 CC therapy procedures. The siRNAs (silencing RNAs) targeted against cellular  
 CC proteins interacting with HIV-1 integrase, LEDGF, MCM7, HBO1, Snurportin,  
 CC VBP1, Transportin-SR and EIF3S3, are useful in preparing a medicament for  
 CC treating HIV-1. The current sequence is that of the human SID protein of  
 CC the invention.

SQ Sequence 302 AA;

Query Match 100.0%; Score 43; DB 7; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9  
 |||||  
 238 NLOQLMDRV 246

RESULT 5

ADBE29059  
 ID ADE29059 standard; protein; 347 AA.

AC ADE29059;

DT 29-JAN-2004 (first entry)

DE Human EIF3S3 selected interacting domain protein - SEQ ID 45.

XX anti-HIV; SID; selected interacting domain; HIV; gene therapy; siRNA;  
 XX silencing RNA; HIV-1 integrase; LEDGF; MCM7; HBO1; Snurportin; VBP1;  
 KW Transportin-SR; EIF3S3; human.

XX Homo sapiens.

XX WO2003046176-A2.

XX PD 05-JUN-2003. X

XX PF 26-NOV-2002; 2002WO-EP013868.

XX PR 26-NOV-2001; 2001US-0333346P.

XX PR 31-MAY-2002; 2002US-0385132P.

XX (HYBR-) HYBRIGENICS.

XX PI Legrain P, Rain J, Benarous R, Emilian S, Berlioz-Torrent C;

XX PI Bloc G;

XX DR WPI: 2003-505199/47.

XX DR N-PSDB; ADE29036.

XX PT New complex between two interacting proteins, useful for screening

XX PT molecules that inhibit human immunodeficiency virus or for preparing a

XX PT medicament for treating HIV-1.

XX PS Claim 9; SEQ ID NO 45; 102pp; English.

XX SQ The invention relates to a novel complex between two interacting proteins

CC listed within the specification. The complex of the invention  
 CC demonstrates anti-HIV activity whilst the SID (selected interacting  
 CC domain) and polypeptides may be useful for screening molecules that  
 CC inhibit human immunodeficiency virus (HIV), as well as during gene  
 CC therapy procedures. The siRNAs (silencing RNAs) targeted against cellular  
 CC proteins interacting with HIV-1 integrase, LEDGF, MCM7, HBO1, Snurportin,  
 CC VBP1, Transportin-SR and EIF3S3, are useful in preparing a medicament for  
 CC treating HIV-1. The current sequence is that of the human SID protein of  
 CC the invention.

SQ Sequence 347 AA;

Query Match 100.0%; Score 43; DB 7; Length 347;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOQLMDRV 9  
 |||||  
 237 NLOQLMDRV 245

RESULT 6

ABB08360  
 ID ABB08360 standard; protein; 352 AA.

AC ABB08360;

DT 07-MAY-2002 (first entry) X

DE Human eIF3 amino acid sequence.

XX Human; melanoma antigen eukaryotic initiation factor 3; eIF3;  
 KW ovarian cancer; MHC; cytosolic; immunomodulator; immune effector cell;  
 KW anti-cancer; vaccine.

XX OS Homo sapiens.  
 XX FH Key  
 FT Region 242. .250  
 FT Binding-site /note= "natural epitope of human cancer antigen eIF3"  
 FT Binding-site 242  
 FT Binding-site /note= "HLA-2 binding residue"  
 FT Binding-site 243  
 FT Binding-site /note= "HLA-2 binding residue"  
 FT Binding-site 244. .249  
 FT Binding-site /note= "T-cell receptor (TCR) binding domain"  
 FT Binding-site 250  
 FT /note= "HLA-2 binding residue"  
 XX FN WO200192307-A2.  
 XX PD 06-DEC-2001.  
 XX PF 30-MAY-2001; 2001WO-US017456.  
 XX PR 31-MAY-2000; 2000US-0209391P.  
 XX PR 17-AUG-2000; 2000US-0226258P.  
 XX PR 20-DEC-2000; 2000US-0257008P.  
 XX (GENZ ) GENZYME CORP.  
 XX PI Nicolette CA;  
 XX DR WPI; 2002-139606/18.  
 XX DR N-PSDB; ABA97211.  
 XX PT New therapeutic compounds useful against human ovarian cancer, for  
 PT modulating immune response in a subject, and for generating antibodies  
 PT that specifically recognize and bind to these molecules.  
 XX PS Disclosure; Page 63-64; 68pp; English.  
 XX CC The invention relates to novel therapeutic compounds, that are designed  
 CC to enhance binding to MHC molecules and to enhance immunoregulatory  
 CC properties relative to their natural counterparts. The activity of the  
 CC compounds of the invention may be described as cytostatic and  
 CC immunomodulatory. The compounds are useful against human ovarian cancer,  
 CC for modulating immune response in a subject, and for generating  
 CC antibodies that specifically recognize and bind to these molecules.  
 CC Compositions comprising the compounds are useful as components of anti-  
 CC cancer vaccines and to expand immune effector cells that are specific for  
 CC cells characterised by expression of antigen eIF3 (melanoma antigen  
 CC eukaryotic initiation factor). The peptides or polypeptides conjugated to  
 CC a detectable agent may be used in diagnostic procedures, such as in the  
 CC detection and purification of antibodies, and as immunogens for  
 CC production of antibodies. The polynucleotides can be used as primers for  
 CC detecting genes or gene transcripts expressed in APC to confirm  
 CC transduction of the polynucleotides into host cells. The current sequence  
 CC represents the human melanoma antigen eukaryotic initiation factor 3  
 CC (eIF3) amino acid sequence  
 XX SQ Sequence 352 AA;  
 QY Query Match 100.0%; Score 43; DB 5; Length 352;  
 DB Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLQQLMDRV 9  
 DB 242 NLQQLMDRV 250  
 RESULT 7  
 ABR82212 standard; protein; 352 AA.  
 AC ABR82212;  
 AC ABR82212;

XX DT 13-OCT-2003 (first entry)  
 XX DE Human antigen eukaryotic translation initiation factor 3 (eIF3).  
 XX KM Eukaryotic translation initiation factor 3; eIF3; neoplasia; cancer;  
 XX KM cytostatic; gene therapy; human; antigen.  
 XX OS Homo sapiens.  
 XX PN WO2003050543-A1.  
 XX PD 19-JUN-2003. X  
 XX PF 05-DEC-2001; 2001WO-US047997.  
 XX PR 05-DEC-2001; 2001WO-US047997.  
 XX (GENZ ) GENZYME CORP.  
 XX PI Nicolette CA;  
 XX DR WPI; 2003-532936/50.  
 XX DR N-PSDB; ACC85029.  
 XX PT Aiding in the diagnosis of a neoplastic condition, useful for treating  
 PT cancer and related malignancies comprises determining the amount of  
 PT expression of an eIF3 protein in a test sample isolated from the cell or  
 PT tissue.  
 XX PS Claim 6; Page 71-72; 77pp; English.  
 XX CC The invention relates to aiding in the diagnosis of a neoplastic  
 CC condition or susceptibility to a neoplastic condition of an animal cell  
 CC or tissue. The method involves determining the amount of expression of an  
 CC eukaryotic translation initiation factor 3 (eIF3) protein in a test  
 CC sample isolated from the cell or tissue, and diagnosing a neoplastic  
 CC condition or susceptibility to a neoplastic condition based on the amount  
 CC of expression of the eIF3 protein. The methods, compounds and kits are  
 CC useful in therapeutics, diagnostic and screening methods for human cancer  
 CC and related malignancies, e.g. ovarian, breast, lung, colon, prostate,  
 CC pancreatic or gastrointestinal cancer, or melanoma. The present sequence  
 CC represents the human antigen eIF3  
 XX SQ Sequence 352 AA;  
 QY Query Match 100.0%; Score 43; DB 7; Length 352;  
 DB Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLQQLMDRV 9  
 DB 242 NLQQLMDRV 250  
 RESULT 8  
 ADJ68541 standard; protein; 352 AA.  
 AC ADJ68541;  
 AC ADJ68541;  
 DT 06-MAY-2004 (first entry)  
 DE Human heat mitochondrial protein as a therapeutic target SEQID347.  
 XX mitochondrial; human; screening assay; diabetes mellitus;  
 XX Huntington's disease; osteoarthritis;  
 XX Leber's hereditary optic neuropathy; LHON;  
 XX mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
 XX myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
 XX osteopathic; ophthalmological; cytostatic.



OS Homo sapiens.  
 XX PN WO2003087768-A2.  
 XX PD 23-OCT-2003.  
 XX PF 04-APR-2003; 2003WO-US010870.  
 XX PR 12-APR-2002; 2002US-0372843P.  
 XX PR 17-JUN-2002; 2002US-0389877P.  
 XX PR 20-SEP-2002; 2002US-0412418P.  
 XX PA (MITO-) MITOKOR.  
 XX PA (BUCK-) BUCK INST AGE RES.  
 XX PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM,  
 XX PI Warnock DE;  
 XX DR WPI; 2003-845369/78.  
 XX PT Identifying a mitochondrial target for drug screening assays and for  
 XX PT treating diseases associated with altered mitochondrial function,  
 XX PT comprises detecting a modified polypeptide in a sample and correlating  
 XX PT with the disease.  
 XX PS Claim 1, SEQ ID NO 347; 180pp; English.  
 XX CC This invention relates to novel mitochondrial targets that can be used  
 XX CC for therapeutic intervention in treating a disease associated with  
 XX CC altered mitochondrial function. Specifically, it refers to a method for  
 XX CC identifying proteins of the human heart mitochondrial proteome that are  
 XX CC useful for drug screening assays, as well as therapeutic targets. The  
 XX CC present invention describes a method for identifying such proteins that  
 XX CC can be used in the treatment of various diseases associated with altered  
 XX CC mitochondrial function including diabetes mellitus, Huntington's disease,  
 XX CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 XX CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 XX CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 XX CC compositions have neuroprotective, nocotropic, antidiabetic,  
 XX CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 XX CC cyostatic activities. This polypeptide sequence is a human heart  
 XX CC mitochondrial protein of the invention.  
 XX SQ Sequence 352 AA;  
 XX  
 XX Query Match 100.0%; Score 43; DB 7; Length 352;  
 XX Best Local Similarity 100.0%; Pred. No. 1.9;  
 XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 XX QY 1 NQLQIMDRV 9  
 XX DB 242 NQLQIMDRV 250  
 XX  
 XX RESULT 9  
 XX ADC31312  
 XX ID ADC31312 standard; protein; 368 AA.  
 XX AC ADC31312;  
 XX XX  
 XX DT 18-DEC-2003 (first entry)  
 XX DE Human novel polypeptide sequence, SEQ ID NO:1394.  
 XX XX  
 XX XX Human; diagnostic; drug screening; forensics; gene mapping;  
 XX XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;  
 XX XX neurodegenerative diseases; anaemia; platelet disorder; wound; burns;  
 XX XX ulcers; osteoporosis; autoimmune disease; cancer;  
 XX XX molecular weight marker; food supplement; antiparkinsonian; nocotropic;  
 XX XX neuroprotective; antiischemic; anticoagulant; thrombolytic; vlnarary;  
 XX XX antifurcer; osteopathic; immunosuppressive; antiinflammatory; cyostatic;  
 XX XX gene therapy; chromosome 6p22.

OS Homo sapiens.  
 XX PN WO2003029271-A2.  
 XX PD 10-APR-2003  
 XX PF 24-SEP-2002; 2002WO-US030474.  
 XX PR 24-SEP-2001; 2001US-0324631P.  
 XX PR (HYSE-) HYSEQ INC.  
 XX XX  
 XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;  
 XX PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;  
 XX PI Haley-Vicente D, Drmanac RT;  
 XX DR WPI; 2003-371981/35.  
 XX DR N-PsDB; ADC30341.  
 XX XX  
 XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or  
 XX PT treating conditions such as neurodegenerative diseases, anemias, platelet  
 XX PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
 XX PT cancer.  
 XX PS Claim 20; SEQ ID NO 1394; 1185pp; English.  
 XX CC The invention relates to 971 novel human cDNA sequences (ADC29919-  
 XX CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The  
 XX CC invention also relates to nucleic acid sequences over 99% identical with  
 XX CC the novel human cDNAs. The invention additionally encompasses expression  
 XX CC vectors and host cells comprising a nucleic acid of the invention; the  
 XX CC recombinant production of a polypeptide of the invention; an antibody  
 XX CC against a polypeptide of the invention; a method of detecting  
 XX CC polynucleotides or polypeptides of the invention; and methods of  
 XX CC identifying a compound which binds to a polypeptide of the invention. The  
 XX CC invention further discloses methods of preventing, treating or  
 XX CC ameliorating a medical condition; kits comprising polynucleotide probes  
 XX CC and/or monoclonal antibodies for carrying out the methods of the  
 XX CC invention; methods for the identification of compounds that modulate the  
 XX CC expression or activity of the polynucleotide and/or polypeptide; and 767  
 XX CC contig sequences corresponding to the cDNA sequences of the invention  
 XX CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628  
 XX CC -ADC33394). The nucleic acids and polypeptides of the invention are  
 XX CC useful in diagnostics, drug screening, forensics, gene mapping, in the  
 XX CC identification of mutations responsible for genetic disorders or other  
 XX CC traits, for assessing biodiversity, and in producing many other types of  
 XX CC data and products dependent on DNA and amino acid sequences. They are  
 XX CC also used for treating diseases such as Parkinson's disease, Alzheimer's  
 XX CC disease and other neurodegenerative diseases, anaemia, platelet  
 XX CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or  
 XX CC cancer. The nucleic acids may also be used as hybridisation probes or  
 XX CC primers, and in the recombinant production of a protein. The polypeptides  
 XX CC are also useful in generating antibodies, as molecular weight markers,  
 XX CC and as food supplements. The present sequence represents a specifically  
 XX CC claimed human polypeptide sequence of the invention. Note: The sequence  
 XX CC data for this patent did not form part of the printed specification, but  
 XX CC was obtained in electronic format directly from WIPO at  
 XX CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX SQ Sequence 368 AA;  
 XX  
 XX Query Match 100.0%; Score 43; DB 7; Length 368;  
 XX Best Local Similarity 100.0%; Pred. No. 2;  
 XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 XX QY 1 NQLQIMDRV 9  
 XX DB 258 NQLQIMDRV 266  
 XX  
 XX RESULT 10  
 XX AAG57646  
 XX ID AAG57646 standard; protein; 260 AA.

XX AAG57646;  
AC 18-OCT-2000 (first entry)  
DT Arabidopsis thaliana protein fragment SEQ ID NO: 74312.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EPI033405-A2.  
XX  
XX  
XX 06-SEP-2000.  
XX  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123160P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125768P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130449P.  
XX 23-APR-1999; 99US-0130510P.  
XX 23-APR-1999; 99US-0130891P.  
XX 28-APR-1999; 99US-0131449P.  
XX 30-APR-1999; 99US-0132048P.  
XX 30-APR-1999; 99US-0132407P.  
XX 04-MAY-1999; 99US-0132484P.  
XX 05-MAY-1999; 99US-0132485P.  
XX 06-MAY-1999; 99US-0132486P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132863P.  
XX 11-MAY-1999; 99US-0134256P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134219P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134768P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136392P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137222P.  
XX 03-JUN-1999; 99US-0137528P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138647P.  
XX 14-JUN-1999; 99US-0138119P.  
XX 16-JUN-1999; 99US-0139452P.  
XX 16-JUN-1999; 99US-0139453P.  
XX 17-JUN-1999; 99US-0139492P.  
XX 18-JUN-1999; 99US-0139454P.  
XX 18-JUN-1999; 99US-0139455P.  
XX 18-JUN-1999; 99US-0139456P.  
XX 18-JUN-1999; 99US-0139457P.  
XX 18-JUN-1999; 99US-0139458P.  
XX 18-JUN-1999; 99US-0139459P.  
XX 18-JUN-1999; 99US-0139460P.  
XX 18-JUN-1999; 99US-0139461P.  
  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148655P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149922P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.

PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150886P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 22-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 83.7%; Score 36; DB 3; Length 260;  
Best Local Similarity 87.5%; Pred. No. 38;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LQLMDRV 9  
Db 95 LQLMDRV1 102

RESULT 11  
AAG57645  
ID AAG57645 standard; protein; 311 AA.  
AC AAG57645;  
XX 18-OCT-2000 (first entry)  
DT Arabidopsis thaliana protein fragment SEQ ID NO: 74311.  
DE  
XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
PD  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130443P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.

PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 26-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 26-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 83.7%; Score 36; DB 3; Length 311;  
Best Local Similarity 87.5%; Pred. No. 46;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQLLMDRV 9  
Db 146 LQLLMDRI 153

RESULT 12  
AAGS7644  
ID AAGS7644 standard; protein; 334 AA.

XX AAGS7644;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 74310.

XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridization assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.  
PD  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-01223180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138647P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147003P.  
PR 04-AUG-1999; 99US-0147044P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 05-AUG-1999; 99US-0147266P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147933P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 25-AUG-1999; 99US-0150884P.  
PR 26-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153756P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.

PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157665P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159285P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 83.7%; Score 36; DB 3; Length 334;  
Best Local Similarity 87.5%; Pred. No. 50;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQIMDRV 9  
DB 169 LQIMDR 176

RESULT 13  
AAB92665  
ID AAB92665 standard; protein; 133 AA.  
XX  
AC AAB92665;  
XX  
DT 26-JUN-2001 (first entry)  
DX  
DE Human protein sequence SEQ ID NO:11024.  
XX  
KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-00116126.  
XX  
PR 29-JUL-1999; 99JP-00248036.  
PR 27-AUG-1999; 99JP-00300253.  
PR 11-JAN-2000; 2000JP-00118776.  
PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
DR WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
XX  
PS Claim 8; SEQ ID NO 11024; 2537pp + Sequence Listing; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 complementary sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the polynucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention  
CC  
CC Sequence 133 AA;

Query Match 81.4%; Score 35; DB 4; Length 133;  
Best Local Similarity 87.5%; Pred. No. 30;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQIMDR 8  
DB 23 NLQIMDR 30

RESULT 14  
AAG01149  
ID AAG01149 standard; protein; 152 AA.  
XX  
AC AAG01149;  
XX  
DT 06-OCT-2000 (first entry)  
DX  
DE Human secreted protein, SEQ ID NO: 5230.  
XX  
KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-00200610.  
XX  
PR 26-FEB-1999; 99US-0122487P.  
XX

PA (GSET) GENSET.  
 XX  
 PI Dumas Milne Edwards J, Duclet A, Giordano J;  
 DR MPI; 2000-500381/45.  
 XX N-PSDB; AAC01155.  
 XX  
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 PS Claim 13; SEQ ID NO 5230; 71bp + Sequence listing; English.  
 XX  
 CC The present sequence is a polypeptide encoded by one of a large number of  
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were  
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different  
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated  
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT  
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA  
 CC sequences derived from the 5' ends of mRNAs and even in those cases where  
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely  
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can  
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs  
 CC are also used in diagnostic, forensic, gene therapy and chromosome  
 CC mapping procedures. They are used to obtain upstream regulatory sequences  
 CC and to design expression and secretion vectors  
 CC  
 SQ Sequence 152 AA;  
 XX  
 XX  
 Query Match 81.4%; Score 35; DB 3; Length 152;  
 Best Local Similarity 87.5%; Pred. No. 34;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLOQLMDR 8  
 |||||:  
 23 NLOQLVDR 30  
 DB  
 RESULT 15  
 AAY28922  
 ID AAY28922 standard; protein; 180 AA.  
 AC AAY28922;  
 XX  
 DT 21-SEP-1999 (first entry)  
 XX  
 DE Human regulatory protein HRGP-8.  
 XX  
 KW Regulatory protein; HRGP; human; cell proliferation; immune response;  
 KW cancer; adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma;  
 KW tetracarcinoma; Addison's disease; adult respiratory distress syndrome;  
 KW allergy; anemia; asthma; atherosclerosis; bronchitis; cholecystitis;  
 KW Crohn's disease; ulcerative colitis; atopic dermatitis; dermatomyositis;  
 KW diabetes mellitus; emphysema; atrophic gastritis; glomerulonephritis;  
 KW gout; Grave's disease; hyperinsinophililia; irritable bowel syndrome; AIDS;  
 KW lupus erythematosus; multiple sclerosis; myasthenia gravis; infection;  
 KW osteoarthritis; osteoporosis; pancreatitis; polymyositis; hemodialysis;  
 KW arthritis; scleroderma; Sjogren's syndrome; and autoimmune thyroiditis;  
 KW infection; truma.  
 KW  
 OS Homo sapiens.  
 XX  
 PN WO933870-A2.  
 XX  
 PD 08-JUL-1999.  
 XX  
 PF 22-DEC-1998; 98WO-US027471.  
 XX  
 PR 31-DEC-1997; 97US-00001403.  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 XX  
 PI Lal P, Bandman O, Hillman JL, Au-Young J, Tang YT, Yue H;

PI Shah P, Guegler KJ, Corley NC;  
 XX  
 DR MPI; 1999-430229/36.  
 XX N-PSDB; AAX89292.  
 XX  
 PT New human regulatory proteins, useful for diagnosing, preventing and  
 PT treating disorders associated with expression of regulatory proteins.  
 XX  
 PS Claim 1; Page 77; 89pp; English.  
 XX  
 CC The invention provides novel human regulatory proteins (AAY28915-926),  
 CC designated HRGP, and their polynucleotides (AAX89285-296). The proteins  
 CC can be produced using standard recombinant technology. The expression of  
 CC HRGP is closely associated with cell proliferation and the polypeptides  
 CC and polynucleotides are useful in the diagnosis, treatment and prevention  
 CC of diseases associated with cell proliferation, particularly immune  
 CC responses and cancer. The protein or agonists may be administered to  
 CC treat or prevent a cancer such as adenocarcinoma, leukemia, lymphoma,  
 CC melanoma, myeloma, sarcoma, and tetracarcinoma. Such cancers include, but  
 CC are not limited to, cancers of the adrenal gland, bladder, bone, brain,  
 CC cervix, breast, gall bladder, ganglia, ovary and pancreas. Where HRGP is  
 CC promoting leukocyte activity or proliferation, antagonists which decrease  
 CC the activity of HRGP are administered. Such responses may be associated  
 CC with disorders such as Addison's disease, adult respiratory distress  
 CC syndrome, AIDS, allergies, anemia, asthma, atherosclerosis, bronchitis,  
 CC cholecystitis, Crohn's disease, ulcerative colitis, atopic dermatitis,  
 CC dermatomyositis, diabetes mellitus, emphysema, atrophic gastritis, gout,  
 CC glomerulonephritis, Grave's disease, hyperinsinophililia, irritable bowel  
 CC syndrome, lupus erythematosus, multiple sclerosis, myasthenia gravis,  
 CC myocardial or pericardial infarction, osteoarthritis, osteoporosis,  
 CC pancreatitis, polyomyositis, rheumatoid arthritis, scleroderma, Sjogren's  
 CC syndrome, and autoimmune thyroiditis, complications of cancer,  
 CC hemodialysis, extracorporeal circulation; viral, bacterial, fungal,  
 CC parasitic, protozoan, and helminthic infections; and trauma. The HRGP  
 CC polynucleotide may be used for diagnosis of these conditions, and as a  
 CC source of primers and probes  
 CC  
 SQ Sequence 180 AA;  
 XX  
 XX  
 Query Match 81.4%; Score 35; DB 2; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 41;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLOQLMDR 8  
 |||||:  
 23 NLOQLVDR 30  
 DB  
 RESULT 16  
 ABB89507  
 ID ABB89507 standard; protein; 180 AA.  
 AC ABB89507;  
 XX  
 DT 24-MAY-2002 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 1883.  
 XX  
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antifungal;  
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.  
 KW  
 OS Homo sapiens.  
 XX  
 PN WO200190304-A2.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 18-MAY-2001; 2001WO-US016450.  
 XX  
 PR 19-MAY-2000; 2000US-0205515P.  
 XX  
 PI

XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Birse CE, Rosen CA;  
 XX MPI: 2002-122018/16.  
 DR N-PSDB; ABI89916.  
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
 PT prevention of neural, immune system, muscular, reproductive,  
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
 disorders.  
 PS Claim 11; SEQ ID NO 1883; 2081pp + Sequence listing; English.  
 XX The invention relates to novel genes (ABL89449-ABL90853) and proteins  
 CC (AB899040-AB899444) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
 CC infections diseases such as viral, bacterial, fungal and parasitic  
 CC infections. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 SQ Sequence 180 AA:  
 QY Query Match 81.4%; Score 35; DB 5; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 41;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 NLOQLMDR 8  
 23 NLOQLVDR 30  
 RESULT 17  
 ID ADD22442 standard; protein; 180 AA.  
 XX ADD22442;  
 AC 15-JAN-2004 (first entry)  
 DT 15-JAN-2004 (first entry)  
 XX HLA-B\*6 T cell recognised tumour antigenic polypeptide, SEQ No 92.  
 DE tumour antigenic peptide; cancer; vaccine; cytotoxic T cell;  
 KW colon; mouth; lung; prostatic; gynecological; human.  
 XX Homo sapiens.  
 OS JP2003111595-A.  
 PN 15-APR-2003.  
 PD 24-JUN-2002; 2002JP-00183603.  
 PF 25-JUN-2001; 2001JP-00191974.  
 PR (ITOY/). ITO Y.  
 PA MPI, 2003-611129/58.  
 XX Novel tumor antigenic peptide or polypeptide useful for inducing  
 PT cytotoxic T cells or for treating cancer such as colon, mouth, lung,

PT prostatic or gynecological cancer.  
 XX Claim 2; SEQ ID NO 92; 98pp; Japanese.  
 PS The invention relates to a novel tumour antigenic peptide or polypeptide  
 CC comprising a sequence selected from 99 sequences fully defined in the  
 CC specification. The tumour antigenic peptide or polypeptide comprises a  
 CC sequence selected from 99 sequences fully defined in the specification.  
 CC where the tumour antigenic peptide preferably has a sequence of Glu-Pro-  
 CC Pro-Leu-Ser-Glu-Glu-Thr-Phe, and the polypeptide preferably has a  
 CC sequence comprising 93 amino acids fully defined in the specification.  
 CC The invention further provides a cancer vaccine comprising a tumour  
 CC antigenic peptide or polypeptide, which has cytostatic activity. A tumour  
 CC antigenic peptide, polypeptide, a recombinant vector containing the  
 CC hybridising polynucleotide, a host transformed with the vector or an antibody are  
 CC useful for screening for compounds that interact with the tumour  
 CC antigenic peptide, the polypeptide or its encoding polynucleotide and  
 CC increases the expression of the tumour antigenic peptide or the polypeptide is  
 CC or polynucleotide. The tumour antigenic peptide or the polypeptide is  
 CC useful for inducing cytotoxic T cells. The tumour antigenic peptide  
 CC vaccine is useful for treating cancer such as colon, mouth, lung,  
 CC prostatic or gynecological cancer. The invention also provides a  
 CC pharmaceutical composition useful for treating cancer. The tumour  
 CC antigenic peptide or the polypeptide is useful as an antigen to create  
 CC antibodies. This sequence represents one of the tumour antigenic  
 CC polypeptides of the invention.  
 SQ Sequence 180 AA:  
 QY Query Match 81.4%; Score 35; DB 7; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 41;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 NLOQLMDR 8  
 23 NLOQLVDR 30  
 RESULT 18  
 ADI15876 standard; protein; 180 AA.  
 ID ADI15876;  
 AC 22-APR-2004 (first entry)  
 DT Human PP 78.  
 XX tumour; antigen; human leukocyte antigen; HLA-A2; HLA-A26;  
 KW cytotoxic T-cell; CTL; vaccine; cancer; colorectal cancer;  
 KW stomach cancer; buccal cancer; renal cancer; lung cancer;  
 KW gynecological cancer; prostate cancer.  
 XX Homo sapiens.  
 OS WO2003008450-A1.  
 PN 30-JAN-2003.  
 PD 11-JUN-2002; 2002WO-JP005799.  
 PF 12-JUN-2001; 2001JP-00177058.  
 PR 21-AUG-2001; 2001JP-00250728.  
 PA (ITOH/) ITOH K.  
 XX Itoh K, Shichiyo S;  
 PI MPI, 2003-267996/26.  
 DR N-PSDB; ADI15951.  
 XX Tumor antigen peptides recognized by human leukocyte antigen (HLA)-A2 or



PT HLA-A26 restricted cytotoxic T-cells for treatment and prevention of  
 PT cancer including preparation of cancer vaccines.  
 PS Claim 2; SEQ ID NO 220; 323bp; Japanese.  
 XX  
 CC The invention relates to a tumour antigen peptide recognised by human  
 CC leukocyte antigen (HLA)-A2 or HLA-A26 restricted cytotoxic T-cells (CTL)  
 CC and/or capable of inducing CTL. The tumour antigen peptide is useful for  
 CC the treatment, prevention, diagnosis and vaccine production for cancers  
 CC including colorectal, stomach, buccal, renal, lung, gynecological and  
 CC prostate cancer. The present sequence represents the amino acid sequence  
 CC of a human protein.  
 XX  
 SQ Sequence 180 AA;  
 Query Match 81.4%; Score 35; DB 7; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 41;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLOQLMDR 8  
 Db 23 NLOQLVDR 30  
 RESULT 19  
 ADI15933  
 ID ADI15933 standard; protein; 180 AA.  
 XX  
 AC ADI15933;  
 XX  
 DT 22-APR-2004 (first entry)  
 XX  
 DE Human PP 76.  
 XX  
 KW tumour; antigen; human leukocyte antigen; HLA-A2; HLA-A26;  
 KW cytotoxic T-cell; CTL; vaccine; cancer; colorectal cancer;  
 KW stomach cancer; buccal cancer; renal cancer; lung cancer;  
 KW gynecological cancer; prostate cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003008450-A1.  
 PD 30-JAN-2003.  
 XX  
 PF 11-JUN-2002; 2002WO-JP005799.  
 XX  
 PR 12-JUN-2001; 2001JP-00177058.  
 PR 21-AUG-2001; 2001JP-00250728.  
 XX  
 PA (ITOY/) ITOH K.  
 XX  
 PI Itoh K, Shichijo S;  
 XX  
 DR WPI; 2003-267996/26.  
 DR N-PSDB; ADI16002.  
 PT  
 PT Tumour antigen peptides recognized by human leukocyte antigen (HLA)-A2 or  
 PT HLA-A26 restricted cytotoxic T-cells for treatment and prevention of  
 PT cancer including preparation of cancer vaccines.  
 XX  
 PS Claim 2; SEQ ID NO 277; 323bp; Japanese.  
 XX  
 CC The invention relates to a tumour antigen peptide recognised by human  
 CC leukocyte antigen (HLA)-A2 or HLA-A26 restricted cytotoxic T-cells (CTL)  
 CC and/or capable of inducing CTL. The tumour antigen peptide is useful for  
 CC the treatment, prevention, diagnosis and vaccine production for cancers  
 CC including colorectal, stomach, buccal, renal, lung, gynecological and  
 CC prostate cancer. The present sequence represents the amino acid sequence  
 CC of a human protein.  
 XX  
 SQ Sequence 180 AA;

Query Match 81.4%; Score 35; DB 7; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 41;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLOQLMDR 8  
 Db 23 NLOQLVDR 30  
 RESULT 20  
 ADJ81699  
 ID ADJ81699 standard; protein; 180 AA.  
 XX  
 AC ADJ81699;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Tumour antigen peptide-related human protein Seqid12.  
 XX  
 KW tumour antigen peptide; cytotoxic T cell; HLA-A constraint property;  
 KW HLA-A3 super type; cytostatic; vaccine; gene therapy; cancer;  
 KW prostatic cancer; colon cancer; stomach cancer; cervical carcinoma;  
 KW breast cancer; lung cancer; oesophageal cancer; bladder cancer; melanoma;  
 KW cancer recognition; HLA-A restricted sexual-cell cytotoxic T cell;  
 KW human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN JP2004000216-A.  
 PD 08-JAN-2004.  
 XX  
 PF 28-APR-2003; 2003JP-00124482.  
 XX  
 PR 26-APR-2002; 2002JP-00126764.  
 XX  
 PA (ITOY/) ITO T.  
 XX  
 DR WPI; 2004-085221/09.  
 DR N-PSDB; ADJ81705.  
 PT  
 PT Novel tumor antigen peptide inducing and/or activating cytotoxic T cells,  
 PT useful as a vaccine for cancer.  
 XX  
 PS Claim 2; SEQ ID NO 12; 77bp; Japanese.  
 XX  
 CC This invention relates to a novel tumour antigen peptide inducing a  
 CC cytotoxic T cell to the HLA-A constraint property which belongs to one or  
 CC more HLA-A3 super type and which attains and is recognised by cytotoxic T  
 CC cells. The invention may be useful for the development of compounds with  
 CC a cytostatic activity or a vaccine. In addition, the sequences disclosed  
 CC may be useful for gene therapy. The invention is useful for treating  
 CC cancer such as prostatic cancer, colon cancer, stomach cancer, cervical  
 CC carcinoma, breast cancer, lung cancer, oesophageal cancer, bladder cancer  
 CC or melanoma. In addition, the invention may be useful for screening  
 CC compounds which increases recognition by HLA-A restricted sexual-cell  
 CC cytotoxic T cells. The invention is useful for fundamental research of  
 CC the molecule related to recognition of the cancer by cytotoxic T cells.  
 CC The invention or compositions developed through its use may be  
 CC efficiently useful in treating cancer and screening compounds which  
 CC increases recognition by HLA-A restricted sexual-cell cytotoxic T cells.  
 CC The present sequence is that of a human protein which was used during the  
 CC derivation of the novel tumour antigen peptides of the invention.  
 XX  
 SQ Sequence 180 AA;  
 Query Match 81.4%; Score 35; DB 8; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 41;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLOQLMDR 8  
 Db 23 NLOQLVDR 30

RESULT 21

AB56651 standard; protein; 239 AA.

AC AB56651;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1229.

Human; prostate cancer; prostate cancer antigen; detection; diagnosis; neuroprotective; cytosolic; cardioprotective; immunomodulatory; muscular; vulnary; gastrointestinal; nephrotoxic; antineoplastic; gynaecological; antibacterial; gene therapy; neural; immune; reproductive; renal; gastrointestinal; pulmonary; cardiovascular; proliferative disorder; wound; infectious disease.

OS Homo sapiens.

PN WO200055174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US005988.

PR 12-MAR-1999; 99US-0124270P.

PA (HUMA-) HUMAN GENOME SCI INC. (ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM;

DR WPI; 2000-587513/55. N-PSDB; AAF15854.

PT Prostate cancer associated gene sequences, referred to as prostate cancer PT antigens, useful for treatment, prevention, and diagnosis of disorders such as prostate cancer.

PS Claim 11; Page 1651-1652; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated CC proteins, called prostate cancer antigens, given in AAB56653 to AAB57302. CC The prostate cancer antigens can have neuroprotective, cytosolic, CC cardioprotective, immunomodulatory, muscular, vulnary, gastrointestinal, CC nephrotoxic, antineoplastic, gynaecological and antibacterial activities, CC and can be used in gene therapy. The prostate cancer antigen CC polynucleotides may be used for detection of prostate cancer, chromosome CC identification, as chromosome markers, and for numerous other diagnostic CC or research purposes. The prostate cancer antigens may be used to treat CC disorders such as neural, immune, muscular, reproductive, CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to CC AAF17303 represent sequences used in the exemplification of the present CC invention

XX Sequence 239 AA;

Query Match 81.4%; Score 35; DB 3; Length 239; Best Local Similarity 87.5%; Pred. No. 55; Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLOQLMDR 8  
DB 82 NLOQLVDR 89

RESULT 22  
ADE55015  
ID ADE55015 standard; protein; 1809 AA.

AC ADE55015;

DT 29-JAN-2004 (first entry)

DE Rat Protein P49816, SEQ ID NO 820.

Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury; chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

PA (GENO ) GEN HOSPITAL CORP. (PARB ) BAYER AG.

PI Woolf C, D'Urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26. GENBANK; P49816.

PT New composition comprising two or more isolated polypeptides, useful for PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat CC or human polynucleotides or a polynucleotide which represents a fragment, CC derivative or allelic variation of the nucleic acid sequence. Also CC claimed are a vector comprising the novel polynucleotide, a host cell CC comprising the vector, a method for identifying a nucleotide sequence CC which is differentially regulated in an animal subjected to pain and a CC kit to perform the method, an array, a method for identifying an agent CC that increases or decreases the expression of the polynucleotide sequence CC that is differentially expressed in neuronal tissue of a first animal CC subjected to pain, a method for identifying a compound which regulates CC the expression of a polynucleotide sequence which is differentially CC expressed in an animal subjected to pain, a method for identifying a CC compound that regulates the activity of one or more of the CC polynucleotides, a method for producing a pharmaceutical composition, a CC method for identifying a compound or small molecule that regulates the CC activity in an animal of one or more of the polypeptides given in the CC specification, a method for identifying a compound useful in creating CC pain and a pharmaceutical composition comprising the one or more CC polypeptides or their antibodies. The polynucleotide or the compound that CC modulates its activity is useful for preparing a medicament for treating CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene CC therapy). The sequence presented is a rat protein (shown in Table 2 of CC the specification) which is differentially expressed during pain. Note: CC The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic form directly from WIPO at CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1809 AA;

Query Match 81.4%; Score 35; DB 7; Length 1809; Best Local Similarity 87.5%; Pred. No. 4.9e+02; Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLOQLMDR 8  
DB 444 NLOQLMER 451

## RESULT 23

AD55009  
ID ADE55009 standard; protein; 1809 AA.

AC ADE55009;

DT 29-JAN-2004 (first entry)

DE Rat Protein P49816, SEQ ID NO 814.

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GENO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

PI WPI, 2003-268312/26.

DR GENBANK; P49816.

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1809 AA;

Query Match 81.4%; Score 35; DB 7; Length 1809;

Best Local Similarity 87.5%; Pred. No. 4.9e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NQLIMDR 8

DB 444 NQLIMDR 451

## RESULT 24

AD55006  
ID ADE55006 standard; protein; 1809 AA.

AC ADE55006;

DT 29-JAN-2004 (first entry)

DE Rat Protein P49816, SEQ ID NO 811.

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GENO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

PI WPI, 2003-268312/26.

DR GENBANK; P49816.

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1809 AA;

Query Match 81.4%; Score 35; DB 7; Length 1809;

Best Local Similarity 87.5%; Pred. No. 4.9e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLOQLMDR 8  
|||:|  
Db 444 NLOQLMER 451

RESULT 25  
ADE55012  
ID ADE55012 standard; protein; 1809 AA.

AC ADE55012;

DT 29-JAN-2004 (first entry)

DE Rat Protein P49816, SEQ ID NO 817.

KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI, 2003-268312/26.

DR GENBANK; P49816.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX SQ Sequence 1809 AA;

Query Match 81.4%; Score 35; DB 7; Length 1809;  
Best Local Similarity 87.5%; Pred. No. 4.9e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLOQLMDR 8  
|||:|  
Db 444 NLOQLMER 451

RESULT 26  
ABB59885  
ID ABB59885 standard; protein; 1116 AA.

AC ABB59885;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 6447.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE ) PE CORP NY.

PA Venter JC, Adams M, Li FWD, Myers EW;

PI WPI, 2001-656860/75.

DR N-PSDB; AB03988.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX SQ Sequence 1116 AA;

Query Match 79.1%; Score 34; DB 4; Length 1116;  
Best Local Similarity 66.7%; Pred. No. 4.7e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NLOQLMDR 9  
|||:|  
Db 630 NLOQLMDR 638

RESULT 27  
ADM26506  
ID ADM26506 standard; protein; 1116 AA.

AC ADM26506;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 6447.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PN 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE ) PE CORP NY.

PA Venter JC, Adams M, Li FWD, Myers EW;

PI WPI, 2001-656860/75.

DR N-PSDB; AB03988.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

ID ADM26506 standard; protein; 509 AA.  
 XX  
 AC ADM26506;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Hyperthermophile Methanopyrus kandleri protein #1112.  
 XX  
 KM hyperthermophile; protein stability enhancement;  
 KW protein activity enhancement.  
 XX  
 OS Methanopyrus kandleri.  
 XX  
 PN WO2003076575-A2.  
 XX  
 PD 18-SEP-2003.  
 XX  
 PF 04-MAR-2003; 2003WO-US006664.  
 XX  
 PR 04-MAR-2003; 2002US-0361742P.  
 PR 14-MAY-2002; 2002US-0380423P.  
 PR 16-SEP-2002; 2002US-0410974P.  
 XX  
 PA (FIDE-) FIDELITY SYSTEMS INC.  
 PA (MALY/) MALYKH A.  
 XX  
 PI Slesarev AI, Pavlov A, Pavlova N, Kozyavkin S;  
 XX  
 DR WPI; 2003-748383/70.  
 DR N-PSDB; ADM27081.  
 XX  
 PT New isolated nucleic acids encoding any of about 1700 Methanopyrus  
 PT kandleri proteins, and the encoded proteins, useful as a medicaments or  
 PT as diagnostic agents.  
 XX  
 PS Claim 31; SEQ ID NO 1112; 1023pp; English.  
 XX  
 CC The invention comprises the amino acid sequence of proteins from the  
 CC hyperthermophile Methanopyrus kandleri, the invention also comprises the  
 CC complete genome from Methanopyrus kandleri. The Methanopyrus kandleri  
 CC proteins of the invention are useful for enhancing the stability and/or  
 CC activity of other proteins. The Methanopyrus kandleri genome is useful in  
 CC a variety of diagnostic and analytical methods. The present amino acid  
 CC sequence represents a Methanopyrus kandleri protein of the invention.  
 CC  
 SQ Sequence 509 AA;  
 XX  
 Query Match 76.7%; Score 33; DB 7; Length 509;  
 Best Local Similarity 77.8%; Pred. No. 3.2e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 NQQLMDRV 9  
 DB 112 NQQLMDRV 120  
 XX  
 RESULT 28  
 AAB39967  
 ID AAB39967 standard; protein; 874 AA.  
 XX  
 AC AAB39967;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Vigna unguiculata starch synthase isoform V protein.  
 XX  
 KM Starch synthase isoform V protein; starch biosynthesis; transgenic plant;  
 KM transgenic; enzyme.  
 XX  
 OS Vigna unguiculata.  
 XX  
 PN US2003097688-A1.  
 XX

PD 22-MAY-2003.  
 XX  
 PF 05-JUN-2002; 2002US-00163214.  
 XX  
 PR 08-JUN-2001; 2001US-0297099P.  
 XX  
 PA (ALLE/) ALLEN S M.  
 PA (BROG/) BROGLIE K E.  
 PA (BUTL/) BUTLER K H.  
 PA (THOR/) THORPE C J.  
 XX  
 PI Allen SM, Broglie KE, Butler KH, Thorpe CJ;  
 XX  
 DR WPI; 2003-765535/72.  
 XX  
 PT Novel isolated starch synthase isoform V polypeptide having starch  
 PT synthase activity, useful for altering starch biosynthesis pathway in  
 PT plants.  
 XX  
 PS Example 3; Fig 1; 40pp; English.  
 XX  
 CC The invention relates to novel starch synthase isoform V proteins having  
 CC starch synthase activity and polynucleotides encoding them. Sequences of  
 CC the invention are useful for transforming a cell. They are useful for  
 CC producing a plant. They are useful for altering starch biosynthesis  
 CC pathway in plants. The present sequence is Vigna unguiculata starch  
 CC synthase isoform V protein  
 CC  
 SQ Sequence 874 AA;  
 XX  
 Query Match 76.7%; Score 33; DB 7; Length 874;  
 Best Local Similarity 75.0%; Pred. No. 5.7e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NQQLMDR 8  
 DB 195 NQQLMDR 202  
 XX  
 RESULT 29  
 ABB63841  
 ID ABB63841 standard; protein; 125 AA.  
 XX  
 AC ABB63841;  
 XX  
 DT 04-NOV-2002 (first entry)  
 XX  
 DE Human ORF211.  
 XX  
 KM Cytostatic; Cardiac; Anti-allergic; Immunosuppressive; Vascular;  
 KM Anti-inflammatory; gene therapy; human; ORF; atherogenic; platelet;  
 KM human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;  
 KM cancer; cardiovascular disease; allergy; autoimmune disease;  
 KM wound healing; blood coagulation disorder; inflammatory disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002082206-A1.  
 XX  
 PD 27-JUN-2002.  
 XX  
 PF 30-MAY-2001; 2001US-00867550.  
 XX  
 PR 30-MAY-2000; 2000US-0208427P.  
 XX  
 PA (LEACH/) LEACH M D.  
 PA (MEHR/) MEHRABAN F.  
 PA (CONL/) CONLEY P B.  
 PA (TOPP/) TOPPER J N.  
 PA (LAWD/) LAW D.  
 XX  
 PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;  
 XX

DR WPI: 2002-626554/67.  
 DR N-PSDB; ABQ98404.  
 XX  
 PT New polypeptide designated ORFX are present in human atherogenic cells  
 PT and are useful to prevent and treat ORFX-associated disorders including  
 PT cancer, allergy, wound healing or autoimmune, cardiovascular or  
 PT inflammatory disease.  
 PS Claim 10; SEQ ID NO 422; 78bp; English.  
 XX  
 CC The present invention relates to novel human ORFX polypeptides and their  
 CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences  
 CC were discovered in human atherogenic cells, in particular in platelets  
 CC and human umbilical vein endothelial cells (HUVEC) and are expressed in  
 CC many other tissues as well. Atherogenic cells are cells which have the  
 CC potential to develop atherosclerotic plaques. The ORFX polypeptides and  
 CC nucleic acids are useful for treating or preventing a pathological  
 CC condition associated with an ORFX-associated disorder, e.g. cancer,  
 CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood  
 CC coagulation disorders or inflammatory disorders. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from the USPTO web site at  
 CC seqdata.uspto.gov/sequence.html?docid=20020082206  
 CC  
 SQ Sequence 125 AA;  
 XX  
 Query Match 74.4%; Score 32; DB 5; Length 125;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLOQLMDR 8  
 DB 22 NLOQLMDQ 29  
 XX  
 RESULT 30  
 AAU44957  
 ID AAU44957 standard; protein; 228 AA.  
 XX  
 AC AAU44957;  
 XX  
 DT 27-FEB-2002 (first entry)  
 XX  
 DE Propionibacterium acnes immunogenic protein #5853.  
 XX  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 OS Propionibacterium acnes.  
 OS  
 PN WO200181581-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 PF 20-APR-2001; 2001WO-US012865.  
 XX  
 PR 21-APR-2000; 2000US-0199047P.  
 PR 02-JUN-2000; 2000US-0208841P.  
 PR 07-JUL-2000; 2000US-0216747P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 DR WPI: 2001-616774/71.  
 XX  
 DR N-PSDB; AAS59524.  
 XX  
 PT Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris.

XX  
 PS Example 1; SEQ ID NO 6152; 1069bp; English.  
 XX  
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for  
 CC this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 SQ Sequence 228 AA;  
 XX  
 Query Match 74.4%; Score 32; DB 4; Length 228;  
 Best Local Similarity 75.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLOQLMDR 8  
 DB 125 NLOQLMDQ 132  
 XX  
 RESULT 31  
 ABM41476  
 ID ABM41476 standard; protein; 228 AA.  
 XX  
 AC ABM41476;  
 XX  
 DT 20-OCT-2003 (first entry)  
 XX  
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #6152.  
 XX  
 KW Acne vulgaris; anti-seborrheic; dermatological; antibacterial;  
 KW immunostimulant; immune response; vaccine.  
 XX  
 OS Propionibacterium acnes.  
 OS  
 PN WO2003033515-A1.  
 XX  
 PD 24-APR-2003.  
 XX  
 PF 11-OCT-2002; 2002WO-US032727.  
 XX  
 PR 15-OCT-2001; 2001US-00978825.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;  
 PI Zhang Y, Wang S, Jen S, Lodes M, Benson DR, Jones R, Carter D;  
 PI Barch B, Valliee-Douglass J;  
 DR WPI: 2003-381789/36.  
 XX  
 DR N-PSDB; ACP64453.  
 XX  
 PT New Propionibacterium acnes polypeptides and polynucleotides encoding the  
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,  
 PT or for stimulating an immune response specific for a P. acnes protein.  
 XX  
 PS Example 1; SEQ ID NO 6152; 1481bp; English.

The invention relates to an isolated polynucleotide (ACF64435-ACF64733) encoding a *Propionibacterium acnes* protein. The invention also relates to polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to immunogenic fragments of *P. acnes* polypeptides. The invention additionally encompasses expression vectors and host cells comprising a polynucleotide of the invention; antibodies against polypeptides of the invention; fusion proteins comprising a polypeptide of the invention; a method for stimulating an immune response specific for a *P. acnes* polypeptide and an isolated T cell population comprising T cells prepared via this method; a vaccine composition (comprising *P. acnes* polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of *P. acnes* in a patient; and a method for inhibiting the development of *P. acnes* in a patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating *acne vulgaris*, or for stimulating an immune response specific for a *P. acnes* protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridization. The vaccine composition is useful for the stimulation of an immune response against *P. acnes*, or for treating *acne*, and the kit is useful for performing a diagnostic assay. The present sequence represents a polypeptide predicted to be encoded by an ORF (open reading frame) contained within the *P. acnes* polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

Query Match	74.4%	Score 32	DB 6	Length 228
Best Local Similarity	75.0%	Pred. No. 2.2e+02		
Matches 6	Conservative	2	Mismatches 0	Indels 0
			Gaps	0

```

QY      1 NLOQLMDR 8
         |||||:|:
Db      125 NLOQLLDQ 132

```

RESULT 32  
ABP28324  
ID ABP28324 standard; protein; 269 AA.

... AC ABP28324;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 5824.

KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

**Streptococcus pyogenes.**

PN WO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB004789.

PR 27-OCT-2000; 2000GB-00026333.

PR 07-MAR-2001; 2001GB-00005640.

PA (CHIR-) CHIRON SPA.

XX

PI Telford J, Massignani V,

PI Tettein H;

[illegible]

DR WPI; 2002-352536/38.

DK N-ESDB; ABN688955.

XX New Streptococcus protein for the treatment or prevention of infection or  
PT disease caused by Streptococcus bacteria, such as meningitis, and for  
PT detecting a compound that binds to the protein.

PS Claim 1; Page 3745; 4525pp; English.

The invention relates to a protein (ABP25413-ABP30895) from group B streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and anti-inflammatory activity. (1), nucleic acids encoding (1), ABN6504-ABN71526 and CC antibodies that bind (1) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*. Nucleic acids encoding (1) are used to detect Streptococcus in a biological sample. (1) is used to determine whether a compound binds to (1). A composition comprising (1) or a nucleic acid encoding (1), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (1) may be used to recombinantly produce (1) and may be used in gene therapy. Antibodies to (1) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins

Query Match	74.4%	Score 32	DB 5	Length 269
Best Local Similarity	85.7%	Pred. NO.	2.6e+02	
Matches 6	Conservative	1	Mismatches 0	Indels 0
			Gaps	0

QY	3	QLMDRV	9
		:	
Db	55	QLMDRI	61

RESULT 33  
ADK46631  
ID ADK46631 standard; protein; 283 AA.

AC ADK46631;

DT 20-MAY-2004 (first entry)

Streptococcus pneumoniae protein, Seq ID No 3146.

**KW** Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.

Streptococcus pneumoniae.

PN US6699703-B1.

PD 02-MAR-2004.

PF 26-MAY-2000; 2000US-00583110.

PR 02-JUL-1997; 97US-0051553P.

PR 30-JUN-1998; 98US-00107433.

PA (GENO-) GENOME THERAPEUTICS C

PI Doucette-Stamm L, Bush D, Zeng

DR WPI; 2004-212399/20.



**PT** preventing and treating

PT screening.	0	4	4	4	4	4

PS Disclosure; SEQ ID NO 3146; 301pp; English.

•

XX The invention relates to isolated Streptococcus pneumoniae nucleic acids  
CC and polypeptides. The nucleic acids and proteins are useful for  
CC diagnosing, preventing and treating pathological conditions resulting  
CC from bacterial infection, such as S. pneumoniae infection. These may also  
CC be used for drug screening procedures. The present sequence represents a  
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
CC data for this patent did not appear in the printed specification but was  
CC obtained in electronic format directly from USPTO at  
CC [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
CC  
CC  
SQ Sequence 283 AA;

Query Match 74.4%; Score 32; DB 8; Length 283;  
Best Local Similarity 66.7%; Pred. No. 2.7e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 NLOQLMDRV 9  
Db 172 NLOQLMDRV 180

RESULT 34  
ABP28470  
ID ABP28470 standard; protein; 621 AA.

AC ABP28470;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 6116.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial;

XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus agalactiae.

PN WO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB004789.

PR 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Telford J, Massignani V, Margarit Y Rosl, Grandi G, Fraser C;

PI Tetteijn H;

DR WPI: 2002-352536/38.

DR N-PSDB: ABN69101.

XX New Streptococcus protein for the treatment or prevention of infection or

PT disease caused by Streptococcus bacteria, such as meningitis, and for

PT detecting a compound that binds to the protein.

XX Claim 1; Page 3774; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and  
CC antibodies that bind (1) are used for the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (1) are used to detect Streptococcus in a  
CC biological sample. (1) is used to determine whether a compound binds to

CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (1) may be used to recombinantly produce (1) and may be  
CC used in gene therapy. Antibodies to (1) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins  
CC  
CC  
SQ Sequence 621 AA;

Query Match 74.4%; Score 32; DB 5; Length 621;  
Best Local Similarity 55.6%; Pred. No. 6.4e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 NLOQLMDRV 9  
Db 155 NLOQLMDRV 163

RESULT 35  
AAO26550  
ID AAO26550 standard; protein; 2697 AA.

AC AAO26550;

DT 06-MAR-2003 (first entry)

DE Cochliobolus ORF protein sequence SEQ ID No 5.

XX Fungicide; antiinflammatory; cyrostatic; library; plant; fungal; avian;

KW mammalian DNA; Cochliobolus; anti-fungal; fungicide; pathogenic;

KW endocrine; gastrointestinal; cardiovascular disorder; cancer; secretory.

XX Cochliobolus sp.

PN US2002142324-A1.

PD 03-OCT-2002.

PF 24-SEP-2001; 2001US-00961527.

PR 22-SEP-2000; 2000US-0234650P.

PR 22-SEP-2000; 2000US-0234673P.

XX (WANG/) WANG X.

PA (TURG/) TURGEON B G.

PA (YODE/) YODER O.

PA (WUJ/) WU J.

PI Wang X, Turgeon BG, Yoder O, Wu J;

PI WPI: 2003-102510/09.

DR N-PSDB: AAL54019.

XX Preparing library of modified DNA fragments involves generating a library

PT of DNA fragments having a double strand break and inserting detectable

PT polynucleotide into the break to yield library of modified DNA fragments.

XX Claim 13; Page 39-45; 65pp; English.

CC The invention relates to a novel method for preparing a library of  
CC modified DNA fragments. The novel method comprises contacting a library  
CC of DNA fragments in a vector with an agent to cause at least one double  
CC strand break in at least one fragment to yield a library of DNA fragments  
CC having at least one double strand break, and inserting a detectable  
CC polynucleotide or gene into the break to yield a library of modified DNA  
CC fragments. The novel method is useful for preparing a cDNA or genomic  
CC library of modified DNA fragments, where the DNA is plant, fungal, avian  
CC or mammalian DNA. The library is useful for identifying the function of a  
CC gene, by contacting cells (plant, bacterial, fungal, avian or mammalian  
CC cells) with the library to yield a population of cells containing at  
CC least one recombinant cell, in which homologous recombination has  
CC occurred between the genome of the cell and the modified DNA in at least



CC one member of the library and identifying the recombinant cell by a  
CC change in phenotype. The isolated Cochliobolus polypeptide of the  
CC invention is useful for identifying an agent having anti-fungal activity.  
CC The anti-fungal agent of the invention is useful as a fungicide to  
CC suppress the growth of pathogenic fungi. The inhibitors of the isolated  
CC Cochliobolus polypeptide may specifically inhibit fungal pathogenicity or  
CC growth and are also useful as a therapeutic in disorders associated with  
CC protein processing and maturation including endocrine, gastrointestinal,  
CC and cardiovascular disorders, in inflammation and in cancers,  
CC particularly those involving secretory and gastrointestinal tissues. This  
CC sequence represents a protein of an ORF region of the Cochliobolus DNA  
CC sequence of the invention  
XX

SQ Sequence 2697 AA;

Query Match 74.4%; Score 32; DB 6; Length 2697;  
Best Local Similarity 75.0%; Pred. No. 3.1e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DY 2 LQLMDRV 9  
DB 605 LQLMDRI 612

RESULT 36

AD042289 standard; protein; 4080 AA.

XX AD042289;

XX 15-JUL-2004 (first entry)

XX Human NOVX polypeptide #69.

XX Human NOVX polypeptide #69.  
XX Human; NOVX; cancer; atherosclerosis; diabetes; Alzheimer's disease;  
XX Parkinson's disease; graft-versus-host disease; scleroderma;  
XX hypertension; haemophilia; idiopathic thrombocytopenic purpura;  
XX immunodeficiency; AIDS; dyslipidemia; obesity; Crohn's disease;  
XX bronchial asthma; anorexia; cancer-associated cachexia;  
XX multiple sclerosis; fertility.

OS Homo sapiens.

XX US2004058338-A1.

XX 25-MAR-2004.

PF 02-DEC-2002; 2002US-00307817.

XX 03-DEC-2001; 2001US-0336881P.  
XX 05-DEC-2001; 2001US-0336820P.  
XX 07-DEC-2001; 2001US-0338285P.  
XX 07-DEC-2001; 2001US-0338318P.  
XX 10-DEC-2001; 2001US-0338989P.  
XX 10-DEC-2001; 2001US-0339022P.  
XX 11-DEC-2001; 2001US-0339314P.  
XX 11-DEC-2001; 2001US-0339516P.  
XX 11-DEC-2001; 2001US-0339517P.  
XX 11-DEC-2001; 2001US-0339611P.  
XX 12-DEC-2001; 2001US-0340981P.  
XX 12-DEC-2001; 2001US-0341346P.  
XX 14-DEC-2001; 2001US-0340390P.  
XX 14-DEC-2001; 2001US-0340440P.  
XX 14-DEC-2001; 2001US-0340565P.  
XX 14-DEC-2001; 2001US-0340608P.  
XX 14-DEC-2001; 2001US-0341144P.  
XX 17-DEC-2001; 2001US-0341477P.  
XX 17-DEC-2001; 2001US-0341540P.  
XX 18-DEC-2001; 2001US-0341768P.  
XX 20-DEC-2001; 2001US-0342529P.  
XX 31-DEC-2001; 2001US-0344903P.  
XX 01-FEB-2002; 2002US-0353286P.  
XX 01-FEB-2002; 2002US-0353288P.

PR 26-FEB-2002; 2002US-0359599P.  
PR 26-FEB-2002; 2002US-0359626P.  
PR 26-FEB-2002; 2002US-0359671P.  
PR 27-FEB-2002; 2002US-0359914P.  
PR 27-FEB-2002; 2002US-0359956P.  
PR 28-FEB-2002; 2002US-0360924P.  
PR 28-FEB-2002; 2002US-0360944P.  
PR 28-FEB-2002; 2002US-0361028P.  
PR 28-FEB-2002; 2002US-0361256P.  
PR 28-FEB-2002; 2002US-0361264P.  
PR 05-MAR-2002; 2002US-0361770P.  
PR 05-MAR-2002; 2002US-0362230P.  
PR 13-MAR-2002; 2002US-0364181P.  
PR 13-MAR-2002; 2002US-0364238P.  
PR 15-MAR-2002; 2002US-0364978P.  
PR 15-MAR-2002; 2002US-0365025P.  
PR 17-APR-2002; 2002US-0373288P.  
PR 15-MAY-2002; 2002US-0380981P.  
PR 16-MAY-2002; 2002US-0381004P.  
PR 17-MAY-2002; 2002US-0381495P.  
PR 28-MAY-2002; 2002US-0383534P.  
PR 28-MAY-2002; 2002US-0383744P.  
PR 29-MAY-2002; 2002US-0383829P.  
PR 29-MAY-2002; 2002US-0384024P.  
PR 02-JUL-2002; 2002US-0393332P.  
PR 06-AUG-2002; 2002US-0401315P.  
PR 07-AUG-2002; 2002US-0401315P.  
PR 20-AUG-2002; 2002US-0404765P.  
PR 23-AUG-2002; 2002US-0405400P.  
PR 23-AUG-2002; 2002US-0405684P.  
PR 23-AUG-2002; 2002US-0405687P.  
PR 23-AUG-2002; 2002US-0405698P.  
PR 26-AUG-2002; 2002US-0406353P.

XX (AGEE/) AGE M. L.  
PA (ALSO/) AUSEROOK J. P.  
PA (ANDE/) ANDERSON D. W.  
PA (BERG/) BERGS L.  
PA (BOLD/) BOLDOG F. L.  
PA (BURG/) BURGESS C. E.  
PA (CATT/) CATTERTON E.  
PA (DIPI/) DIPIRO V. A.  
PA (EDIN/) EDINGER S. R.  
PA (EISE/) EISEN A.  
PA (GANG/) GANGLIOLI E. A.  
PA (GERL/) GERLACH V.  
PA (GORM/) GORMAN L.  
PA (GOTH/) ROTHBERG B. G.  
PA (GUOX/) GUO X. S.  
PA (HERR/) HERRMANN J. L.  
PA (HALV/) HALVORSEN Y.  
PA (UTW/) UT W.  
PA (KERU/) KERUDA R.  
PA (KHRA/) KHRAMTSOV N. V.  
PA (LARO/) LAROCHELLE W. J.  
PA (LEPL/) LEPLLEY D. M.  
PA (LILL/) LI L.  
PA (MACD/) MACDOUGALL J. R.  
PA (MILL/) MILLER C. E.  
PA (ORTT/) ORT T.  
PA (PADI/) PADIGARU M.  
PA (PATT/) PATTURAJAN M.  
PA (PENR/) PENNA C. E. A.  
PA (PEYM/) PEYMAN J. A.  
PA (RIEG/) RIEGER D. K.  
PA (ROTH/) ROTHENBERG M. E.  
PA (SHEN/) SHENOY S. G.  
PA (SMIT/) SMITHSON G.  
PA (SPAD/) SPADERNA S. K.  
PA (SPYT/) SPYTEK K. A.  
PA (STON/) STONE D. J.  
PA (TAUP/) TAUPIER R. J.

PA (VERN/) VERNET C A M.  
 PA (VOSS/) VOSS E Z.  
 PA (ZHON/) ZHONG M.  
 XX  
 PI Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FI,  
 PI Burgess CE, Carterton E, Dipippo VA, Edinger SR, Eisen A,  
 PI Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG, Guo XS,  
 PI Herrmann UL, Halvorsen Y, Ji W, Kekuda R, Khramtsov NV,  
 PI Larochele WJ, Lepley DM, Li L, MacDougall JR, Miller CE, Ort T,  
 PI Padigaru M, Patutajan M, Pena CE, Peyman JA, Rieger DK,  
 PI Rothenberg ME, Shenoy SG, Smithson G, Spaderna SK, Spytek KA,  
 PI Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;  
 XX  
 DR WPI; 2004-268786/25.  
 DR N-PSDB; AD042288.  
 XX  
 PT New human NOVX polypeptides and nucleic acid molecules, useful for  
 PT diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,  
 PT atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or  
 PT scleroderma.  
 PT  
 PS Claim 1; SEQ ID NO 138; 610pp; English.  
 XX  
 CC The invention relates to human NOVX polypeptides and the polynucleotides  
 CC encoding them. The invention also relates to antibodies specific to the  
 CC NOVX polypeptides. The polypeptides, polynucleotides and antibodies are  
 CC useful for manufacturing a medicament for treating a syndrome associated  
 CC with a human disease, such as a pathology associated with the NOVX  
 CC polypeptide. The sequences are useful for diagnosing, treating or  
 CC preventing a NOVX-associated disorder, e.g., cancer, atherosclerosis,  
 CC diabetes, Alzheimer's disease, Parkinson's disease, graft-versus-host  
 CC disease, scleroderma, hypertension, haemophilia, idiopathic  
 CC thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,  
 CC obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated  
 CC cachexia, multiple sclerosis or fertility. The nucleic acids may be used  
 CC as hybridisation probes, in chromosome mapping, in tissue typing, in  
 CC preventive medicine or in pharmacogenomics. This sequence represents a  
 CC human NOVX polypeptide of the invention.  
 CC  
 SQ Sequence 4080 AA;  
 XX  
 QY Query Match 74.4%; Score 32; DB 8; Length 4080;  
 Best Local Similarity 66.7%; Pred. No. 4.8e+03;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 DB 1 NLOQLMDRV 9  
 796 DLRLDRV 804  
 XX  
 RESULT 37  
 ADE78572  
 ID ADE78572 standard; protein; 4624 AA.  
 XX  
 AC ADE78572;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Dynein axonemal heavy polypeptide 5 (DNH5) protein.  
 XX  
 KW dynein axonemal heavy polypeptide 5; DNH5; cilia dysmotility;  
 KW male infertility; sterility; prenatal diagnosis; ciliary;  
 KW flagellar defect; outer dynein arm; human.  
 XX  
 OS Homo sapiens.  
 OS  
 PN EPI327684-A1.  
 PN  
 XX 16-JUL-2003.  
 PD  
 XX 14-JAN-2002; 2002EP-00000820.  
 PF  
 XX 14-JAN-2002; 2002EP-00000820.  
 PR

XX  
 PA (UNIV-) UNIVERSITAETSKLINIKUM FREIBURG.  
 XX  
 PI Omran H;  
 XX  
 DR WPI; 2003-723305/69.  
 DR N-PSDB; ADE78571.  
 XX  
 PT Novel DNH5 polypeptide of molecular motor dynein present in cilia and  
 PT flagella, useful for diagnosing primary ciliary dyskinesia.  
 PT  
 PS Claim 6; SEQ ID NO 2; 53pp; English.  
 XX  
 CC The invention relates to a novel isolated and purified polypeptide having  
 CC a fully defined dynein axonemal heavy polypeptide 5 (DNH5) sequence of  
 CC 4624 amino acids as given in specification, or its active part, amide,  
 CC ester or salt. The invention further relates to a method for the  
 CC detection of a mutation in the DNH5 polynucleotide is useful for  
 CC detecting cilia dysmotility in an afflicted person. The tandem repeat  
 CC polymorphisms are used for the detection. The DNH5 polynucleotide is  
 CC useful for diagnosis of a disease related to cilia dysmotility by  
 CC detecting the presence or absence of DNH5 polynucleotides in cells of  
 CC the afflicted person. The DNH5 polynucleotide is also useful for  
 CC diagnosis of male infertility or sterility, where a genetic analysis of  
 CC the DNH5 polynucleotide in an afflicted person is carried out. The DNH5  
 CC polynucleotide, its complement/fragment or a vector comprising the DNH5  
 CC polynucleotide, are useful for introducing the DNH5 gene in a human cell  
 CC having a defect in the gene to express or to increase the expression  
 CC level of DNH5 gene. The method of the invention is also useful for  
 CC prenatal diagnosis. An antibody derived from a DNH5 protein is useful  
 CC for diagnosis of PCD, male infertility or sterility, ciliary and  
 CC flagellar defects of the outer dynein arm. The antibody is also useful  
 CC for quantitating DNH5 or its salt, amide or ester. This sequence  
 CC represents the dynein axonemal heavy polypeptide 5 (DNH5) protein of the  
 CC invention.  
 CC  
 SQ Sequence 4624 AA;  
 XX  
 QY Query Match 74.4%; Score 32; DB 7; Length 4624;  
 Best Local Similarity 66.7%; Pred. No. 5.5e+03;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 DB 1 NLOQLMDRV 9  
 812 DLRLDRV 820  
 XX  
 RESULT 38  
 ADJ71185  
 ID ADJ71185 standard; protein; 4624 AA.  
 XX  
 AC ADJ71185;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Human heat mitochondrial protein as a therapeutic target SeqID2991.  
 XX  
 KW mitochondrial; human; screening assay; diabetes mellitus;  
 KW Huntington's disease; osteoarthritis;  
 KW Leber's hereditary optic neuropathy; LHON;  
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;  
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
 KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;  
 KW osteopathic; ophthalmological; cyostatic.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO2003087768-A2.  
 PN  
 XX 23-OCT-2003.  
 PD  
 XX 04-APR-2003; 2003WO-US010870.  
 PF  
 XX

PR 12-APR-2002; 2002US-0372843P.  
 PR 17-JUN-2002; 2002US-0389987P.  
 PR 20-SEP-2002; 2002US-0412418P.  
 XX  
 XX  
 PA (MITO-) MITOKOR.  
 PA (BUCK-) BUCK INST AGE RES.  
 XX  
 PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM,  
 PI Warnock DE;  
 XX  
 DR WPI; 2003-845369/78.  
 XX  
 PT Identifying a mitochondrial target for drug screening assays and for  
 PT treating diseases associated with altered mitochondrial function.  
 PT comprises detecting a modified polypeptide in a sample and correlating  
 PT with the disease.  
 XX  
 PS Claim 1, SEQ ID NO 2991; 180pp; English.  
 XX  
 CC This invention relates to novel mitochondrial targets that can be used  
 CC for therapeutic intervention in treating a disease associated with  
 CC altered mitochondrial function. Specifically, it refers to a method for  
 CC identifying proteins of the human heart mitochondrial proteome that are  
 CC useful for drug screening assays, as well as therapeutic targets. The  
 CC present invention describes a method for identifying such proteins that  
 CC can be used in the treatment of various diseases associated with altered  
 CC mitochondrial function including diabetes mellitus, Huntington's disease,  
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 CC osteophalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 CC compositions have neuroprotective, neurotropic, antidiabetic,  
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 CC cyostatic activities. This polypeptide sequence is a human heart  
 CC mitochondrial protein of the invention.  
 XX  
 SQ Sequence 4624 AA;  
 XX  
 Query Match 74.4%; Score 32; DB 7; Length 4624;  
 Best Local Similarity 66.7%; Pred. No. 5.5e+03;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NQQLMDRV 9  
 Db 812 DUEILDRV 820  
 XX  
 RESULT 39  
 ADK48512  
 ID ADK48512 standard; protein; 152 AA.  
 AC ADK48512;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Streptococcus pneumoniae protein, Seq ID No 5027.  
 XX  
 KM Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.  
 XX  
 OS Streptococcus pneumoniae.  
 XX  
 PN US6699703-B1.  
 XX  
 PD 02-MAR-2004.  
 XX  
 PF 26-MAY-2000; 2000US-00583110.  
 XX  
 PR 02-JUL-1997; 97US-0051553P.  
 PR 12-MAY-1998; 98US-0085131P.  
 PR 30-JUN-1998; 98US-00107433.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Housewart CB;

XX  
 DR WPI; 2004-212399/20.  
 DR N-PSDB; ADK45851.  
 XX  
 XX  
 PT New nucleic acid molecules and polypeptides useful for diagnosing,  
 PT preventing and treating pathological conditions resulting from bacterial  
 PT infection, e.g. Streptococcus pneumoniae infection, and in drug  
 PT screening.  
 XX  
 PS Disclosure; SEQ ID NO 5027; 301pp; English.  
 XX  
 CC The invention relates to isolated Streptococcus pneumoniae nucleic acids  
 CC and polypeptides. The nucleic acids and proteins are useful for  
 CC diagnosing, preventing and treating pathological conditions resulting  
 CC from bacterial infection, such as S. pneumoniae infection. These may also  
 CC be used for drug screening procedures. The present sequence represents a  
 CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence  
 CC data for this patent did not appear in the printed specification but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html.  
 CC  
 SQ Sequence 152 AA;  
 XX  
 Query Match 72.1%; Score 31; DB 8; Length 152;  
 Best Local Similarity 55.6%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 NQQLMDRV 9  
 Db 132 NQQLMDRV 140  
 XX  
 RESULT 40  
 AAB76847  
 ID AAB76847 standard; protein; 270 AA.  
 AC AAB76847;  
 XX  
 DT 11-APR-2001 (first entry)  
 XX  
 DE Corynebacterium glutamicum MCT protein SEQ ID NO:676.  
 XX  
 KM Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;  
 KM membrane construction and membrane transport protein; petroleum sp11;  
 KM hydrocarbon degradation; gram positive aerobic bacterium; marker;  
 KM identification; microorganism; fine chemical production; transformation;  
 KM genome mapping; genetic engineering.  
 XX  
 OS Corynebacterium glutamicum.  
 XX  
 PN WO200100805-A2.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 23-JUN-2000; 2000WO-IB000926.  
 XX  
 PR 25-JUN-1999; 99US-0141031P.  
 PR 08-JUL-1999; 99DE-01031454.  
 PR 08-JUL-1999; 99DE-01031478.  
 PR 08-JUL-1999; 99DE-01031563.  
 PR 09-JUL-1999; 99DE-01032122.  
 PR 09-JUL-1999; 99DE-01032124.  
 PR 09-JUL-1999; 99DE-01032125.  
 PR 09-JUL-1999; 99DE-01032128.  
 PR 09-JUL-1999; 99DE-01032180.  
 PR 09-JUL-1999; 99DE-01032182.  
 PR 09-JUL-1999; 99DE-01032190.  
 PR 09-JUL-1999; 99DE-01032191.  
 PR 09-JUL-1999; 99DE-01032209.  
 PR 09-JUL-1999; 99DE-01032212.  
 PR 09-JUL-1999; 99DE-01032227.  
 PR 09-JUL-1999; 99DE-01032228.  
 PR 09-JUL-1999; 99DE-01032229.

PR 09-JUL-1999; 99DE-01032230.  
 PR 14-JUL-1999; 99DE-01032927.  
 PR 14-JUL-1999; 99DE-01033005.  
 PR 14-JUL-1999; 99DE-01033006.  
 PR 27-AUG-1999; 99DE-01040764.  
 PR 27-AUG-1999; 99DE-01040765.  
 PR 27-AUG-1999; 99DE-01040766.  
 PR 27-AUG-1999; 99DE-01040830.  
 PR 27-AUG-1999; 99DE-01040831.  
 PR 27-AUG-1999; 99DE-01040832.  
 PR 27-AUG-1999; 99DE-01040833.  
 PR 31-AUG-1999; 99DE-01041378.  
 PR 31-AUG-1999; 99DE-01041379.  
 PR 31-AUG-1999; 99DE-01041395.  
 PR 03-SEP-1999; 99DE-01042077.  
 PR 03-SEP-1999; 99DE-01042078.  
 PR 03-SEP-1999; 99DE-01042079.  
 PR 03-SEP-1999; 99DE-01042088.

(BAD1) BASF AG.

Pompejus M, Kroegeer B, Schroeder H, Zelder O, Habernauer G;

WPI; 2001-071466/08.  
 N-PSDB; AAF68080.

Corynebacterium glutamicum nucleic acids encoding membrane construction and membrane transport proteins or their portions, useful for typing or identifying C. glutamicum or related bacteria, and as markers for transformation.

Claim 20; Page 1118-1119; 1119pp; English.

AAf67743 to AAF68080 encode the Corynebacterium glutamicum membrane construction and membrane transport (MCT) proteins given in AAB76510 to AAB76847. The MCT nucleic acids and proteins are useful in the identification of microorganisms which can be used to produce fine chemicals, for modulating fine chemical production in C. glutamicum or related bacteria (e.g. Brevibacterium lactofermentum), the typing or identification of C. glutamicum or related bacteria, as reference points for mapping C. glutamicum genome, and as markers for transformation. AAF68082 and AAF68082 represent sequencing primers which are used in an example from the present invention

SQ Sequence 270 AA;

Query Match 72.1%; Score 31; DB 4; Length 270;

Best Local Similarity 66.7%; Pred. No. 4.2e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQILMDRV 9  
 |||:||||  
 Db 22 NLPLIVDRV 30

Search completed: January 12, 2005, 20:07:08  
 Job time : 205.6 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:46:14 ; Search time 16.4 Seconds  
(without alignments)

52.802 Million cell updates/sec

Title: US-09-870-216c-11

Perfect score: 43

Sequence: 1 NLQIMDRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	83.7	311	2	G86324
2	35	81.4	1809	2	S57329
3	34	79.1	1116	2	T13854
4	32	74.4	283	2	B98065
5	32	74.4	305	2	F69306
6	32	74.4	572	2	S73730
7	32	74.4	1244	2	S73731
8	31	72.1	258	2	B72697
9	31	72.1	270	2	D69791
10	31	72.1	291	2	A82416
11	31	72.1	445	1	S74826
12	31	72.1	467	2	C84420
13	31	72.1	519	2	B84811
14	31	72.1	582	2	A90192
15	31	72.1	651	2	JUN057
16	31	72.1	652	2	AD0546
17	31	72.1	672	2	A72076
18	31	72.1	672	2	B86548
19	31	72.1	841	2	JC0647
20	31	72.1	861	2	E97473
21	31	72.1	861	2	A17691
22	31	72.1	1504	2	A33602
23	30	69.8	129	2	T09329
24	30	69.8	144	2	F91130
25	30	69.8	144	2	P85975
26	30	69.8	217	2	T44404
27	30	69.8	218	2	AE1147
28	30	69.8	219	2	AE1506
29	30	69.8	227	2	A75304

30	30	69.8	244	2	T19750	hypothetical prote
31	30	69.8	252	2	E86339	protein F2D10.16 l
32	30	69.8	269	2	E97726	DNA-directed DNA p
33	30	69.8	272	2	T15820	hypothetical prote
34	30	69.8	275	1	B64077	bia(5'-nucleosyl)-
35	30	69.8	284	2	E69113	hypothetical prote
36	30	69.8	316	2	T01637	gene R protein - A
37	30	69.8	333	2	T23840	hypothetical prote
38	30	69.8	340	2	E71801	probable O-sialogl
39	30	69.8	349	2	D88431	protein M8.4 [imp
40	30	69.8	356	2	H90168	GTP-binding protei
41	30	69.8	361	2	PC4295	aspartate-tRNA lig
42	30	69.8	464	2	T48449	fatty acid elongas
43	30	69.8	483	2	T48328	importin alpha-1k
44	30	69.8	508	2	C82138	conserved hypothet
45	30	69.8	548	2	T41092	hypothetical prote

## ALIGNMENTS

### RESULT 1

G86324

hypothetical protein T29W8.1 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C:Accession: G86324

R:Theologos, A.; Becker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.; ansen, N.F.; Hughes, B.; Huzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: G86324

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <STO>

A:Cross-references: UNIPROT:Q9LMB2; GB:AE005172; NID:98954052; PIDN:AAF82225.1; GSPDB:GN

C:Genetics:

A:Map position: 1

Query Match

Best Local Similarity

Matches

7; Conservative

1; Mismatches

0; Indels

0; Gaps

0;

0;

QY 2 LQIMDRV 9

DB 146 LQIMDRV 153

RESULT 2

S57329

tuberos sclerosis 2 homolog - rat

N:Alternate names: tuberlin

C:Species: Rattus norvegicus (Norway rat)

C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 09-Jul-2004

R:Kobayashi, T.; Nishizawa, M.; Hirayama, Y.; Kobayashi, E.; Hino, O.

Nucleic Acids Res. 23, 2608-2613, 1995

A>Title: cDNA structure, alternative splicing and exon-intron organization of the predi-

A:Reference number: S57329; MUID:95380273; PMID:7651821

A:Accession: S57329

A:Status: preliminary; nucleic acid sequence not shown

83.7%	Score 36;	DB 2;	Length 311;
87.5%	Pred. No. 4.8;		
1;	Mismatches	0;	Indels
0;	Gaps	0;	

R:Xiao, G.  
Cell Growth Differ. 6, 1185-1191, 1995  
A:Title: Identification of tubercous sclerosis 2 messenger RNA splice variants that are c  
A:Reference number: 152834; MUID:96063895; PMID:8519635  
A:Accession: 152834  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-931; 'K', 933-1513, 'C', 1515-1729, 'R', 1731-1809 <RES>  
A:Cross-references: EMBL:U04150; NID:91061324; PIDN:AMC52289.1; PID:91061325

Query Match  
Best Local Similarity 81.4%; Score 35; DB 2; Length 1809;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOQLMDR 8  
DB 444 NLOQLMDR 451

RESULT 3  
T13854  
nuclear protein SA - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: T13854  
R:Valdeolmillos, A.M.; Villares, R.; Buesa, J.M.; Gonzalez-Crespo, S.; Martinez, A.; Bar  
DNA Cell Biol. 8, 699-706, 1998  
A:Title: Molecular cloning and expression of stromalin protein from *Drosophila melanogae*  
A:Reference number: Z17798  
A:Accession: T13854  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1116 <VAL>  
A:Cross-references: UNIPROT:O18415; EMBL:Y14277; NID:e1202704; PID:e330100; PIDN:CAA7465  
A:Genetics:

Query Match  
Best Local Similarity 79.1%; Score 34; DB 2; Length 1116;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 9  
DB 630 NLOQLMDR 638

RESULT 4  
E98065  
conserved hypothetical protein spr1550 [Imported] - *Streptococcus pneumoniae* (strain R6)  
C:Species: *Streptococcus pneumoniae*  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 09-Jul-2004  
C:Accession: E98065  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; Dehoff, B.S.; E  
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnren, S.; N  
Y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.R.;  
A:Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: E98065  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-283 <KUR>  
A:Cross-references: UNIPROT:O8DNT6; GB:AE007317; PIDN:AL00354.1; PID:915459215; GSPDB:C  
C:Genetics:  
A:Gene: spr1550

Query Match  
Best Local Similarity 74.4%; Score 32; DB 2; Length 283;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 9  
DB 111 NLOQLMDR 9

DB 172 NLOQLMDR 180

RESULT 5  
F69306  
hypothetical protein AF0454 - *Archaeoglobus fulgidus*  
C:Species: *Archaeoglobus fulgidus*  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: F69306  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E.F.  
; Glodek, A.; Zhou, L.; Overbeek, R.; Cooney, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Utecht, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: F69306  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-305 <KLE>  
A:Cross-references: UNIPROT:O29795; GB:AB01073; GB:AB00782; NID:92689396; PIDN:AAB9078

Query Match  
Best Local Similarity 74.4%; Score 32; DB 2; Length 305;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 9  
DB 222 NLOQLMDR 230

RESULT 6  
S73730  
MG307 homolog H08 orf572c - *Mycoplasma pneumoniae* (strain ATCC 29342)  
C:Species: *Mycoplasma pneumoniae*  
A:Variety: ATCC 29342  
C>Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S73730  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*  
A:Reference number: S73377; MUID:97105885; PMID:8948633  
A:Accession: S73730  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-572 <HIM>  
A:Cross-references: UNIPROT:P75341; EMBL:AE000039; GB:U00089; NID:91674082; PIDN:AAB9601  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
A:Genetics: SGC3

Query Match  
Best Local Similarity 74.4%; Score 32; DB 2; Length 572;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOQLMDR 9  
DB 428 NLOQLMDR 436

RESULT 7  
S73731  
probable lipoprotein A05 orf1244 - *Mycoplasma pneumoniae* (strain ATCC 29342)  
N:Alternate names: MG307 homolog A05\_orf1244  
C:Species: *Mycoplasma pneumoniae*  
A:Variety: ATCC 29342  
C>Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S73731  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*  
A:Reference number: S73377; MUID:97105885; PMID:8948633

A:Accession: S73731  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1244 <K1M>  
 A:Cross-references: UNIPROT:P75342; EMBL:AF000039; GB:U00089; NID:G1674082; PIDD:AB9605  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
 C:Genetics:  
 A:Genetic code: SGC3  
 C:Superfamily: Mycoplasma hypothetical protein MG309

Query Match 74.4%; Score 32; DB 2; Length 1244;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLOLMMDR 9  
 Db 1082 NLOLMMDR 1090

RESULT 8  
 B72697  
 hypothetical protein APE0994 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C:Accession: B72697

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999  
 A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: B72697

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-258 <KAN>

A:Cross-references: UNIPROT:Q9YDB9; DBJ:AP000060; NID:G5104188; PIDD:BAA79978.1; PIDD:dl

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0994

Query Match 72.1%; Score 31; DB 2; Length 258;  
 Best Local Similarity 75.0%; Pred. No. 48;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQOLMDR 9

Db 224 LRLMDR 231

RESULT 9

D69791  
 conserved hypothetical protein yead - Bacillus subtilis

C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C:Accession: D69791

R:Kunst, F.; Ogatawa, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capano, V.; Carter, N.M.; Cho A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hultio, M.F. Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardiniois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel Y., M.; Ogawa, K.; Ogilvara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeder, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Serot akeshi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Wintere, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zumeit, E.; Yoshikawa, H.; Danchin, A.  
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: D69791

A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA

A:Residues: 1-270 <KUN>  
 A:Cross-references: UNIPROT:P94475; GB:Z99107; GB:AL009126; NID:G2632866; PIDD:CA12453.  
 A:Experimental source: strain 168  
 C:Genetics:

A:Gene: yead

Query Match 72.1%; Score 31; DB 2; Length 270;  
 Best Local Similarity 87.5%; Pred. No. 50;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLOLMMDR 8  
 Db 115 NLOLMMDR 122

RESULT 10  
 AE2416  
 hypothetical protein all4885 (imported) - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120  
 A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
 C:Accession: AE2416

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S DNA Res. 8, 205-213, 2001

A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE2416

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-291 <KUN>

A:Cross-references: UNIPROT:Q8YMP8; GB:BA000019; PIDD:BAF6584.1; PIDD:G17134022; GSPDB:G

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all4885

Query Match 72.1%; Score 31; DB 2; Length 291;  
 Best Local Similarity 75.0%; Pred. No. 55;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLOLMMDR 8

Db 170 NLOLMMDR 177

RESULT 11

S74826  
 NADH2 dehydrogenase (EC 1.6.99.3) - Synechocystis sp. (strain PCC 6803)

N:Alternate names: protein slr0851

C:Species: Synechocystis sp.  
 A:Variety: PCC 6803

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: S74826

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N., o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996  
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S74826

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-445 <KAN>

A:Cross-references: UNIPROT:P73739; EMBL:D90909; GB:AB001339; NID:G1652844; PIDD:BA1778

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: ndh

C:Superfamily: NADH dehydrogenase

C:Keywords: electron transfer; FAD; flavoprotein; NAD; oxidoreductase  
 Query Match 72.1%; Score 31; DB 1; Length 445;  
 Best Local Similarity 55.6%; Pred. No. 88;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOLMDRV 9  
 ||:|||||  
 Db 76 NLRVLMMDRV 84

## RESULT 12

hypothetical protein At2g01070 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cross)  
 C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C/Accession: C84420  
 R/Lin. X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.;  
 Eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A/Reference number: A84420; MUID:20083487; PMID:10617197  
 A/Accession: C84420  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-467 <STO>  
 A/Cross-references: UNIPROT:Q95JY9; GB:AE002093; NID:96598622; PIDN:AAFI8655.1; GSPDB:GN  
 C/Genetics:  
 A/Map position: 2

Query Match 72.1%; Score 31; DB 2; Length 467;  
 Best Local Similarity 62.5%; Pred. No. 93;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOLMDRV 8  
 ||:|||||  
 Db 411 NMRLMDRV 418

## RESULT 13

hypothetical protein At2g38960 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cross)  
 C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C/Accession: E84811  
 R/Lin. X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.;  
 Eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A/Reference number: A84420; MUID:20083487; PMID:10617197  
 A/Accession: E84811  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-519 <STO>  
 A/Cross-references: UNIPROT:Q9ZV11; GB:AE002093; NID:93928083; PIDN:AACT9609.1; GSPDB:GN  
 C/Genetics:  
 A/Map position: 2

Query Match 72.1%; Score 31; DB 2; Length 519;  
 Best Local Similarity 77.8%; Pred. No. 1.1e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLOLMDRV 9  
 ||:|||||  
 Db 267 NLRVLMMDRV 275

## RESULT 14

conserved hypothetical protein [imported] - Sulfolobus solfataricus  
 C/Species: Sulfolobus solfataricus  
 C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
 C/Accession: A90192  
 R/She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.  
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Coet, J.  
 submitted to Genbank, April 2001  
 A/Description: Sulfolobus solfataricus complete genome.  
 A/Reference number: A99139

A/Accession: A90192  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-582 <KUR>  
 A/Cross-references: UNIPROT:Q97ZV6; GB:AE006641; NID:913813624; PIDN:AAK40792.1; GSPDB:GN  
 C/Genetics:  
 A/Map position: 2

Query Match 72.1%; Score 31; DB 2; Length 582;  
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLOLMDRV 9  
 ||:|||||  
 Db 571 NLRVLMMDRV 579

## RESULT 15

methylintransferase (EC 2.1.1.-) - Salmonella typhimurium  
 C/Species: Salmonella typhimurium  
 C/Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 09-Jul-2004  
 C/Accession: JN0657  
 R/Darlot, V.; De Backer, O.; Colson, C.  
 Gene 127, 105-110, 1993  
 A/Title: Sequence of the Salmonella typhimurium SxyL11 restriction-modification genes: H  
 A/Reference number: JN0657; MUID:9325265; PMID:8387444  
 A/Accession: JN0657  
 A/Molecule type: DNA  
 A/Residues: 1-651 <DAR>  
 A/Cross-references: UNIPROT:P40814; GB:M90544  
 A/Comment: This enzyme belongs to the type-III restriction-modification system and is ne  
 C/Genetics:  
 A/Key words: DNA binding; methylintransferase; restriction modification system  
 F/13-137/Region: S-adenosylmethionine binding #status predicted

Query Match 72.1%; Score 31; DB 2; Length 651;  
 Best Local Similarity 55.6%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLOLMDRV 9  
 ||:|||||  
 Db 215 NLRVLMMDRV 223

## RESULT 16

site-specific DNA-methylintransferase (adenine-specific) (EC 2.1.1.72) - Salmonella enter  
 A/Species: Salmonella enterica subsp. enterica serovar typh  
 A/Note: this species has also been called Salmonella typh  
 C/Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C/Accession: AD0546  
 R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar  
 S.; Moule, S.; O'Garra, P.  
 Nature 413, 848-852, 2001  
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
 A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero  
 A/Reference number: AD0546; MUID:21534947; PMID:11677608  
 A/Accession: AD0546  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-652 <PAR>  
 A/Cross-references: GB:AL513382; PIDN:CA008812.1; PID:916501627; GSPDB:GN00176  
 C/Genetics:  
 A/Key words: methylintransferase; S-adenosylmethionine



Query Match 72.1%; Score 31; DB 2; Length 652;  
Best Local Similarity 55.6%; Pred. No. 1.4e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NQOLMDRV 9  
|||:||||:  
Db 216 NQOLMDRI 224

## RESULT 17

hypothetical protein CP0290 [imported] - Chlamydia pneumoniae (strains CML029 and AF

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: A72076; A81594

R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; PMID:9206606; PMID:10192388

A:Accession: A72076  
A:Status: Preliminary  
A:Molecule type: DNA

A:Residues: 1-672 <ARN>  
A:Cross-references: UNIPROT:Q92887; GB:AE001630; GB:AE001363; NID:g4376740; PIDN:ADD1860

A:Experimental source: strain CML029  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; PMID:20150255; PMID:10684935

A:Accession: A81594  
A:Status: Preliminary  
A:Molecule type: DNA

A:Residues: 1-672 <REA>  
A:Cross-references: GB:AE002190; GB:AE002161; NID:g7189209; PIDN:AAF8147.1; PID:g718921

A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0462; CP0290

C:Superfamily: Chlamydia hypothetical protein CP0462

Query Match 72.1%; Score 31; DB 2; Length 672;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NQOLMDRV 9  
|||:||||:  
Db 507 NQOLMDRI 515

## RESULT 18

hypothetical protein CP0462 [imported] - Chlamydia pneumoniae (strain J138)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: B86548

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie  
Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.  
A:Reference number: A86491; PMID:20330349; PMID:10871362

A:Accession: B86548  
A:Status: Preliminary  
A:Molecule type: DNA

A:Residues: 1-672 <STO>  
A:Cross-references: UNIPROT:Q92887; GB:BA000008; NID:g8978832; PIDN:BA08666.1; GSPDB:GN  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: CP0462

C:Superfamily: Chlamydia hypothetical protein CP0462

Query Match 72.1%; Score 31; DB 2; Length 672;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NQOLMDRV 9  
|||:||||:  
Db 507 NQOLMDRI 515

## RESULT 19

preprotein translocase secA - Bacillus subtilis

N:Alternate names: Div protein; secA protein homolog  
C:Species: Bacillus subtilis  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: J00647; S17771; F69704

R:Sadale, Y.; Takamatsu, H.; Nakamura, K.; Yamane, K.  
Gene 98, 101-105, 1991

A:Title: Sequencing reveals similarity of the wild-type div+ gene of Bacillus subtilis t  
A:Reference number: J00647; PMID:91192600; PMID:1901557

A:Accession: J00647  
A:Molecule type: DNA

A:Residues: 1-841 <SAD>  
A:Cross-references: UNIPROT:P28366; GB:D10279; DDBJ:D90218; NID:g216332; PIDN:BA01122.1

A:Experimental source: strain Marburg 168r  
R:Overhoff, B.; Klein, M.; Spies, M.; Frendl, R.  
Mol. Gen. Genet. 228, 417-423, 1991

A:Title: Identification of a gene fragment which codes for the 364 amino-terminal amino  
export apparatus in gram-positive and gram-negative bacteria.  
A:Reference number: S17771; PMID:91375427; PMID:1832735

A:Accession: S17771  
A:Molecule type: DNA

A:Residues: 1-125, 'I', 127-364 <OVS>  
A:Cross-references: EMBL:X62035; NID:g48979; PIDN:CAA3977.1; PID:g48980

R:Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997

A:Authors: Folger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler  
isch, J.; Harwood, C.R.; Henuit, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudga, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; PMID:98044033; PMID:9384377

A:Accession: F69704  
A:Status: Preliminary  
A:Molecule type: DNA

A:Residues: 1-841 <KIN>  
A:Cross-references: GB:Z99122; GB:AL009126; NID:g2636029; PIDN:CAB15547.1; PID:g2636056

A:Experimental source: strain 168  
C:Comment: This pleiotropic protein is required for cell division, sporulation, septatio  
is adjacent to the "nucleotide-binding motif B" and "DEAD motif" features as annotated are c  
C:Genetics:  
A:Gene: secA; div

C:Superfamily: preprotein translocase secA  
C:Keywords: ATP; membrane-associated complex; P-loop; protein transport; sporulation  
F/100-107/Region: nucleotide-binding motif A (P-loop) #status atypical  
F/203-208/Region: nucleotide-binding motif B  
F/207-210/Region: DEXH motif

Query Match 72.1%; Score 31; DB 2; Length 841;  
Best Local Similarity 62.5%; Pred. No. 1.8e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LQOLMDRV 9  
|||:||||:  
Db 685 LQOLMDRI 692

## RESULT 20

E97473  
1708 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)  
C/Species: Agrobacterium tumefaciens  
C/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C/Accession: E97473  
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Molim, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A/Reference number: A97359; MUID:21608551; PMID:11743194  
A/Accession: E97473  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-861 <KUR>  
A/Cross-references: UNIPROT:Q8UGU8; GB:AE007869; PIDN:AAK6742.1; PID:915155936; GSPDB:C  
C/Genetics:  
A/Genes: AGR C.1708  
A/Map position: circular cation-transporting ATPase APE1454; ATPase nucleotide-binding domain  
C/Superfamily: probable cation-transporting ATPase APE1454; ATPase nucleotide-binding domain

Query Match 72.1%; Score 31; DB 2; Length 861;  
Best Local Similarity 75.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQILMDRV 9  
DB 443 IQLVDRV 450

## RESULT 21

A12691  
copper transporting ATPase Atcu0937 [imported] - Agrobacterium tumefaciens (strain C58, D  
C/Species: Agrobacterium tumefaciens  
C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C/Accession: A12691  
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.  
erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; Mclell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A/Reference number: AB2577; MUID:21608550; PMID:11743193  
A/Accession: A12691  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-861 <KUR>  
A/Cross-references: UNIPROT:Q8UGU8; GB:AE008688; PIDN:AAL41951.1; PID:917739319; GSPDB:C  
A/Experimental source: strain C58 (Dupont)  
C/Genetics:  
A/Genes: Atcu0937  
A/Map position: circular chromosome  
C/Superfamily: probable cation-transporting ATPase APE1454; ATPase nucleotide-binding domain

Query Match 72.1%; Score 31; DB 2; Length 861;  
Best Local Similarity 75.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQILMDRV 9  
DB 443 IQLVDRV 450

## RESULT 22

A33602  
DNA-directed DNA polymerase (EC 2.7.7.7) REV3 - Yeast (Saccharomyces cerevisiae)  
N/Alternate names: protein p2535; protein YPL167c  
C/Species: Saccharomyces cerevisiae  
C/Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 09-Jul-2004  
C/Accession: A33602; S65178; S69432  
R/Morrison, A.; Christensen, R.B.; Alley, J.; Beck, A.K.; Bernatine, E.G.; Lemontt, J.F.  
J. Bacteriol. 171, 5659-5667, 1989

A/Title: REV3, a Saccharomyces cerevisiae gene whose function is required for induced mut  
A/Reference number: A33602; MUID:90008808; PMID:2676986  
A/Accession: A33602  
A/Molecule type: DNA  
A/Residues: 1-1504 <MOR>  
A/Cross-references: UNIPROT:P14284; GB:M29683; NID:9172386; PIDN:AAA4968.1; PID:9172387  
A/Experimental source: strain S288C  
R/Purnelle, B.; Coster, F.; Goffeau, A.  
submitted to the Protein Sequence Database, May 1996  
A/Reference number: S65154  
A/Accession: S65178  
A/Molecule type: DNA  
A/Residues: 1-1504 <PUR>  
A/Cross-references: EMBL:Z73523; NID:91370352; PIDN:CAA97873.1; PID:91370353; MIPS:YPL16  
A/Experimental source: strain S288C (AB972)  
R/Purnelle, B.; Coster, F.; Naveau, F.; Goffeau, A.  
submitted to the EMBL Data Library, March 1996  
A/Description: The sequence of 55 kb on the left arm of Yeast chromosome XVI identifies  
ogues to the human phosphotyrosyl phosphatase activator PTPA and a homologue to the plant

A/Reference number: S69428  
A/Accession: S69432  
A/Molecule type: DNA  
A/Residues: 1-1504 <PUR>  
A/Cross-references: EMBL:X86770; NID:91403537; PIDN:CA65554.1; PID:91403542  
C/Genetics:  
A/Genes: SGD:REV3; PS01  
A/Cross-references: SGD:S0006088; MIPS:YPL167c  
A/Map position: 16L  
C/Superfamily: yeast DNA-directed DNA polymerase REV3  
C/Keywords: DNA binding; nucleotidyltransferase

Query Match 72.1%; Score 31; DB 2; Length 1504;  
Best Local Similarity 75.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQILMDR 8  
DB 237 DLQILMDR 244

## RESULT 23

T09329  
X1LF1 protein - human herpesvirus 6 (strain U1102)  
C/Species: human herpesvirus 6  
C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C/Accession: T09329  
R/Nicholas, J.; Martin, M.  
U. Virol. 68, 597-610, 1994  
A/Title: Nucleotide sequence analysis of a 38.5-kilobase-pair region of the genome of h  
A/Reference number: Z16644; MUID:94118404; PMID:8289364  
A/Accession: T09329  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-129 <NIC>  
A/Cross-references: UNIPROT:Q69058; EMBL:L25528; NID:9451932; PIDN:AAA16742.1; PID:9451  
C/Genetics:  
A/Genes: X1LF1

Query Match 69.8%; Score 30; DB 2; Length 129;  
Best Local Similarity 66.7%; Pred. No. 36;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQILMDRV 9  
DB 33 NLELLYERV 41

## RESULT 24

P01130  
hypothetical protein BC64014 [imported] - Escherichia coli (strain O157:H7, substrain R  
C/Species: Escherichia coli  
C/Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C/Accession: P01130

R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. Gasawara, N.; Yasunaga, T.; Kohana, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
 DNA Res. 8, 11-22, 2001  
 A>Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc  
 A:Reference number: A39629; MUID:21156231; PMID:11258796  
 A:Accession: F91130  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-144 <HAY>  
 A:Cross-references: UNIPROT:Q8XAC4; UNIPROT:Q8FDA9; GB:BA000007; PIDN:BA037437.1; PID:91  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 A:Genetics:  
 A:Gene: EC64014

Query Match 69.8%; Score 30; DB 2; Length 144;  
 Best Local Similarity 66.7%; Pred. No. 41;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQILMDRV 9  
 |||||  
 Db 93 NLQILMDRV 101

RESULT 25  
 P85975  
 Hypothetical protein Z4488 [imported] - *Escherichia coli* (strain O157:H7, substrain BDL9  
 C:Species: *Escherichia coli*  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: P85975  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobbeck, E.J.; Davis, N.W.; Lam, A.; Diallanta, E.; Potamousis, K.; Apodaca,  
 Nucleic 409, 523-533, 2001  
 A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: P85975  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-144 <STO>  
 A:Cross-references: UNIPROT:Q8XAC4; UNIPROT:Q8FDA9; GB:AE005174; NID:G12517732; PIDN:AAQ  
 A:Experimental source: strain O157:H7, substrain EDL933  
 A:Genetics:  
 A:Gene: Z4488

Query Match 69.8%; Score 30; DB 2; Length 144;  
 Best Local Similarity 66.7%; Pred. No. 41;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQILMDRV 9  
 |||||  
 Db 93 NLQILMDRV 101

RESULT 26  
 T44404  
 adenylate kinase (EC 2.7.4.3) [imported] - *Bacillus halodurans*  
 C:Species: *Bacillus halodurans*  
 C:Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 09-Jul-2004  
 C:Accession: T44404; C83669  
 R:Takami, H.; Takaki, Y.; Nakasone, K.; Hirama, C.; Inoue, A.; Horikoshi, K.  
 Biotech. Biotechnol. Biochem. 63, 452-455, 1999  
 A>Title: Sequence analysis of a 32-kb region including the major ribosomal protein gene  
 A:Reference number: Z22756; MUID:99209008; PMID:10192928  
 A:Accession: T44404  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-217 <TAK>  
 A:Cross-references: UNIPROT:P38372; EMBL:AB017508; NID:G4512395; PIDN:BA075292.1; PID:94  
 A:Experimental source: strain C-125  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
 A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: C83669

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-217 <STO>  
 A:Cross-references: GB:AP001507; GB:BA000004; NID:G10172612; PIDN:BA03874.1; GSPDB:GN00  
 A:Experimental source: strain C-125  
 A:Genetics:  
 A:Gene: adk  
 C:Superfamily: adenylate kinase  
 C:Keywords: phosphotransferase

Query Match 69.8%; Score 30; DB 2; Length 217;  
 Best Local Similarity 85.7%; Pred. No. 65;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QILMDRV 9  
 |||||  
 Db 118 QILMDRV 124

RESULT 27  
 AE1147  
 carboxylesterase homolog lmo0580 [imported] - *Listeria monocytogenes* (strain EGD-e)  
 C:Species: *Listeria monocytogenes*  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AE1147  
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, U.; Kunz, M.; Kunz, F.; Kurapkai, G.; Madueno, E.; Maltournam, A.; Ma  
 ok, C.; Schluter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
 A>Title: Comparative genomics of *Listeria species*.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AE1147  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-218 <GLA>  
 A:Cross-references: UNIPROT:Q8Y9E9; GB:NC\_003210; PIDN:CAC98659.1; PID:G16409956; GSPDB  
 A:Experimental source: strain EGD-e  
 A:Genetics:  
 A:Gene: lmo0580

Query Match 69.8%; Score 30; DB 2; Length 218;  
 Best Local Similarity 85.7%; Pred. No. 66;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQILMDRV 8  
 |||||  
 Db 38 LQILMDRV 44

RESULT 28  
 AE1506  
 weakly carboxylesterase homolog lln0589 [imported] - *Listeria innocua* (strain C1p11262)  
 C:Species: *Listeria innocua*  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C:Accession: AE1506  
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, U.; Kunz, M.; Kunz, F.; Kurapkai, G.; Madueno, E.; Maltournam, A.; Ma  
 ok, C.; Schluter, T.; Simoes, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,  
 A>Title: Comparative genomics of *Listeria species*.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AE1506  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-219 <GLA>  
 A:Cross-references: UNIPROT:Q92E71; GB:AL592022; PIDN:CAC95821.1; PID:G16413029; GSPDB:G  
 A:Experimental source: strain C1p11262  
 A:Genetics:  
 A:Gene: lln0589

Query Match 69.8%; Score 30; DB 2; Length 219;  
 Best Local Similarity 85.7%; Pred. No. 66;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQIMDR 8  
 |||||  
 Db 38 LQIMDR 44

RESULT 29  
 A75304  
 ABC transporter, ATP-binding protein - Deinococcus radiodurans (strain R1)  
 C/Species: Deinococcus radiodurans  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C/Accession: A75304  
 R. White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A/Reference number: A75250; MUID:20036896; PMID:10567286  
 A/Accession: A75304  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-227 <WHI>  
 A/Cross-references: UNIPROT:Q9RS03; GB:AE002052; GB:AE000513; NID:g6459990; PIDN:AAF1174  
 A/Experimental source: strain R1  
 C/Genetics:  
 A/Gene: DR2192  
 A/Map position: 1  
 C/Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 69.8%; Score 30; DB 2; Length 227;  
 Best Local Similarity 85.7%; Pred. No. 69;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QIMDRV 9  
 |||||  
 Db 125 QIMDRV 131

RESULT 30  
 T19750  
 hypothetical protein C35C5.7 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C/Accession: T19750  
 R. White, S.  
 submitted to the EMBL Data Library, August 1996  
 A/Reference number: Z19173  
 A/Accession: T19750  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-244 <WHI>  
 A/Cross-references: EMBL:Z78417; PIDN:CAB01687.1; GSPDB:GN00028; CESP:C35C5.7  
 A/Experimental source: clone C35C5  
 C/Genetics:  
 A/Gene: CESP:C35C5.7  
 A/Map position: X  
 A/Introns: 51/1; 67/2; 103/3; 150/2; 201/3

Query Match 69.8%; Score 30; DB 2; Length 244;  
 Best Local Similarity 62.5%; Pred. No. 74;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NQIMDR 8  
 |||||  
 Db 166 NQIMDR 173

RESULT 31  
 E86339

protein F2D10.16 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C/Accession: E86339  
 R. Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizer, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luhrs, J.S.; Malt, R.; Matzfeld,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A/Reference number: A86141; MUID:21016719; PMID:11130712  
 A/Accession: E86339  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-252 <STO>  
 A/Cross-references: UNIPROT:Q9LM87; GB:AE005172; NID:g8886927; PIDN:AAF0613.1; GSPDB:GN  
 C/Genetics:  
 A/Gene: F2D10.16  
 A/Map position: 1

Query Match 69.8%; Score 30; DB 2; Length 252;  
 Best Local Similarity 66.7%; Pred. No. 77;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NQIMDRV 9  
 |||||  
 Db 56 NQIMDRV 64

RESULT 32  
 E97726  
 DNA-directed DNA polymerase (EC 2.7.7.7) - Rickettsia conorii (strain Malish 7)  
 C/Species: Rickettsia conorii  
 C/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
 C/Accession: E97726  
 R. Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R  
 Science 293, 2093-2098, 2001  
 A/Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A/Reference number: A97700; MUID:21442074; PMID:11557893  
 A/Accession: E97726  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-269 <KIR>  
 A/Cross-references: UNIPROT:Q92056; GB:AE006914; PIDN:AL02751.1; PID:G15619265; GSPDB:GN  
 C/Genetics:  
 A/Gene: holB  
 C/Keywords: nucleotidyltransferase

Query Match 69.8%; Score 30; DB 2; Length 269;  
 Best Local Similarity 66.7%; Pred. No. 83;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NQIMDRV 9  
 |||||  
 Db 196 NQIMDRV 204

RESULT 33  
 T15820  
 hypothetical protein C52B11.2 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: T15820  
 R. Martin, J.  
 submitted to the EMBL Data Library, November 1995  
 A/Description: The sequence of C. elegans cosmid C52B11.  
 A/Reference number: Z18411  
 A/Accession: T15820  
 A/Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-272 <MAR>  
 A:Cross-references: UNIPROT:Q18776; EMBL:U41276; NID:G1086884; PID:G1086886; PIDN:AAA824  
 A:Experimental source: strain Bristol N2  
 C:Genetics:  
 A:Gene: CESP:CS2B11.2  
 A:Introns: 12/1; 91/3; 190/3; 221/3

Query Match 69.8%; Score 30; DB 2; Length 272;  
 Best Local Similarity 71.4%; Pred. No. 84;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 QLMDRV 9  
 ||:||||  
 Db 174 QLMDRI 180

RESULT 34  
 B64077  
 bis(5'-nucleosyl)-tetraphosphatase (symmetrical) (EC 3.6.1.41) - Haemophilus influenzae  
 N:Alternate names: didenostine tetraphosphatase (symmetrical)  
 C:Species: Haemophilus influenzae  
 C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004  
 C:Accession: B64077  
 R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995  
 A:Author: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95550630; PMID:7542800  
 A:Accession: B64077  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-275 <TIGR>  
 A:Cross-references: UNIPROT:P44751; GB:U32737; GB:L42023; NID:G1573536; PIDN:AA022209.1;  
 A:Experimental source: strain Rd KW20  
 C:Genetics:  
 A:Note: TIGR:HI0551  
 C:Function:  
 A:Pathway: purine metabolism  
 C:Superfamily: bis(5'-nucleosyl)-tetraphosphatase (symmetrical); phosphoesterase core hc  
 C:Keywords: hydrolase  
 F:2-67/Domain: phosphoesterase core homology <PBC>

Query Match 69.8%; Score 30; DB 1; Length 275;  
 Best Local Similarity 75.0%; Pred. No. 85;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LQIMDRV 9  
 ||:||||  
 Db 16 LQIMLERY 23

RESULT 35  
 B69113  
 hypochelical protein MTH1843 - Methanobacterium thermoautotrophicum (strain Delta H)  
 C:Species: Methanobacterium thermoautotrophicum  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
 C:Accession: B69113  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; K.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N. J. Bacteriol. 179, 7135-7155, 1997  
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: B69113  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-284 <MTH>  
 A:Cross-references: UNIPROT:O27871; GB:AB000937; GB:AE000666; NID:G2622974; PIDN:AA8630  
 A:Experimental source: strain Delta H  
 C:Genetics:

A:Gene: MTH1843

Query Match 69.8%; Score 30; DB 2; Length 284;  
 Best Local Similarity 85.7%; Pred. No. 88;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLQIMLD 7  
 ||:||||  
 Db 57 NLRLMD 63

RESULT 36  
 T01637  
 gene R protein - Acinetobacter calcoaceticus (fragment)  
 C:Species: Acinetobacter calcoaceticus  
 C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 22-Jun-1999  
 C:Accession: T01637; G32252  
 R:Goosen, N.; Horsman, H.P.A.; Huinen, R.G.M.; van de Putte, P. J. Bacteriol. 171, 447-455, 1989  
 A:Title: Acinetobacter calcoaceticus genes involved in biosynthesis of the coenzyme pyr.  
 A:Reference number: A32252; MUID:89123056; PMID:2536663  
 A:Accession: T01637  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-316 <GOO>  
 A:Cross-references: EMBL:X06452; NID:G38740; PIDN:CAA29758.1; PID:G38746  
 A:Experimental source: strain LMD 79.41  
 C:Superfamily: membrane dipeptidase

Query Match 69.8%; Score 30; DB 2; Length 316;  
 Best Local Similarity 66.7%; Pred. No. 99;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQIMDRV 9  
 ||:||||  
 Db 281 HLEFLMDRV 289

RESULT 37  
 T23840  
 hypochelical protein M88.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T23840  
 R:Sulston, J. submitted to the EMBL Data Library, June 1994  
 A:Reference number: Z19806  
 A:Accession: T23840  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-333 <WIL>  
 A:Cross-references: UNIPROT:O21602; EMBL:Z34802; PIDN:CAA64335.2; GSPDB:GN00021; CESP:M8  
 A:Experimental source: clone M88  
 C:Genetics:  
 A:Gene: CESP:M88.4  
 A:Map position: 3  
 A:Introns: 31/3; 55/3; 93/2; 140/3; 168/2; 224/3; 279/2

Query Match 69.8%; Score 30; DB 2; Length 333;  
 Best Local Similarity 75.0%; Pred. No. 116+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQIMDR 8  
 ||:||||  
 Db 138 NLQIMLSR 145

RESULT 38  
 E71801  
 probable o-sialoglycoprotein endopeptidase - Helicobacter pylori (strain J99)  
 C:Species: Helicobacter pylori  
 A:Variety: strain J99  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004

C:Accession: E71801  
 R:Alm, R.A.; Gibson, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
 A:Reference number: A71800; MUID:99120557; PMID:9923682  
 A:Accession: E71801  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-340 <ARN>  
 A:Cross-references: UNIPROT:Q9ZJ27; GB:AE001570; GB:AE001439; NID:g4156108; PIDN:AND0706  
 A:Experimental source: strain J99  
 C:Genetics:  
 A:Gene: ydiE  
 C:Superfamily: O-sialoglycoprotein endopeptidase

Query Match 69.8%; Score 30; DB 2; Length 340;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQILMDRV 9  
 |||:|  
 Db 54 NLPLLRV 62

RESULT 39  
 D88431  
 protein M88.4 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
 C:Accession: D88431  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_ele  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: D88431  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-349 <STO>  
 A:Cross-references: UNIPROT:Q21602; GB:chr\_III; PIDN:CAA84335.1; PID:g3878685; GSPDB:GNC  
 C:Genetics:  
 A:Gene: M88.4  
 A:Map position: 3

Query Match 69.8%; Score 30; DB 2; Length 349;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQILMDR 8  
 |||:|  
 Db 154 NLQILMR 161

RESULT 40  
 H90168  
 GTP-binding protein (hflX) [imported] - Sulfolobus solfataricus  
 C:Species: Sulfolobus solfataricus  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
 C:Accession: H90168  
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
 Jong, I.; Jeffries, A.C.; Kozera, C.W.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F  
 arett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
 submitted to GenBank, April 2001  
 A:Description: Sulfolobus solfataricus complete genome.  
 A:Reference number: A99139  
 A:Accession: H90168  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-356 <KUR>  
 A:Cross-references: UNIPROT:Q980W3; GB:AB006641; NID:g13813407; PIDN:AAK40607.1; GSPDB:G  
 C:Genetics:

A:Gene: hflX  
 Query Match 69.8%; Score 30; DB 2; Length 356;  
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLQILMDRV 9  
 |||:|  
 Db 340 NLPLLRDKI 348

Search completed: January 12, 2005, 20:15:50  
 Job time : 18.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2005, 19:38:08 ; Search time 85 Seconds

(without alignments)  
60.922 Million cell updates/sec

Title: US-09-870-216c-11

Perfect score: 43

Sequence: 1 NIGLIMRV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	352	1 IF33_HUMAN	O15372 homo sapien
2	43	100.0	352	2 Q6P908	Q6P908 ractus novy
3	43	100.0	352	3 AAH60586	AAH60586 ractus no
4	43	100.0	352	2 CAG33187	CAG33187 homo sapi
5	38	88.4	335	2 Q6P381	Q6P381 xenopus tro
6	38	88.4	335	2 AAH64151	AAH64151 xenopus t
7	38	88.4	352	1 IF33_MOUSE	IF33_MOUSE
8	38	88.4	352	2 Q8BTX5	Q8BTX5 mus musculu
9	37	86.0	473	2 Q8DJU7	Q8DJU7 synecococc
10	36	83.7	196	2 Q94BU3	Q94BU3 arabidopsis
11	36	83.7	311	2 Q8LAP2	Q8LAP2 arabidopsis
12	36	83.7	311	2 Q9LMB2	Q9LMB2 arabidopsis
13	36	83.7	311	2 Q9LMB3	Q9LMB3 arabidopsis
14	35	81.4	133	2 Q9NZ20	Q9NZ20 homo sapien
15	35	81.4	180	2 Q9Y221	Q9Y221 homo sapien
16	35	81.4	180	2 Q9WV50	Q9WV50 ractus novy
17	35	81.4	180	2 Q9CXK8	Q9CXK8 mus musculu
18	35	81.4	180	2 Q9D1B4	Q9D1B4 mus musculu
19	35	81.4	180	2 AAH59114	AAH59114 ractus no
20	35	81.4	180	2 BAD05056	BAD05056 homo sapi
21	35	81.4	1742	2 Q7TT21	Q7TT21 mus musculu
22	35	81.4	1742	2 AAH60701	AAH60701 mus muscu
23	35	81.4	1809	1 TSC2_RAT	TSC2_RAT
24	35	81.4	1814	1 TSC2_MOUSE	TSC2_MOUSE
25	34	79.1	149	2 Q6T1Z6	Q6T1Z6 ractus novy
26	34	79.1	149	2 AAP20218	AAP20218 pagrus majo
27	34	79.1	285	2 Q31388	Q31388 cyprinus ca
28	34	79.1	560	2 Q73MF8	Q73MF8 treponema d
29	34	79.1	560	2 AAS12067	AAS12067 treponema
30	34	79.1	681	1 RROC_ANTRO	RROC_ANTRO
31	34	79.1	681	2 Q6BZ11	Q6BZ11 anthoceros
32	34	79.1	681	2 Q6BZ11	Q6BZ11 debaryomyce

32	34	79.1	964	2 Q7UUS8	Q7UUS8 rhodopirell
33	34	79.1	1116	2 O18415	O18415 drosophila
34	34	79.1	1127	2 Q9VM62	Q9VM62 drosophila
35	34	79.1	1127	2 AAF52463	AAF52463 drosophila
36	33	76.7	257	2 Q9WBR8	Q9WBR8 staphylococ
37	33	76.7	345	2 Q6F7M9	Q6F7M9 acinetobact
38	33	76.7	509	2 Q8TWG3	Q8TWG3 methanopyru
39	33	76.7	874	2 Q9XGCI	Q9XGCI vigna unguil
40	32	74.4	35	2 Q88G77	Q88G77 pseudomonas
41	32	74.4	99	2 Q7SFM5	Q7SFM5 neurospora
42	32	74.4	99	2 CAE76192	CAE76192 neurospor
43	32	74.4	175	2 Q6ZSS5	Q6ZSS5 homo sapien
44	32	74.4	175	2 BAC6871	BAC6871 homo sapi
45	32	74.4	181	2 Q6BGJ7	Q6BGJ7 paramecium

## ALIGNMENTS

RESULT 1  
ID IF33\_HUMAN STANDARD; PRT; 352 AA.  
AC O15372;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Eukaryotic translation initiation factor 3 subunit 3 (eIF-3 gamma)  
DE (eIF3 p40 subunit) (eIF3h).  
GN Name=EIF3S3;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP TISSUE=Liver;  
RC MEDLINE=98001678; PubMed=9341143;  
RX Asano K., Vornlocher H.-P., Richter-Cook N.J., Merrick W.C.,  
RA Hinnebusch A.G., Hershey J.W.B.,  
RT "Structure of cDNAs encoding human eukaryotic initiation factor 3  
RT subunits. Possible roles in RNA binding and macromolecular assembly.",  
RL U. Biol. Chem. 272:27042-27052(1997).  
[2]  
SEQUENCE FROM N.A.  
RX Schmidt O.G., von Holtum D., Gross S., Horsthemke B., Lueddecke H.-J.,  
RA "The gene encoding the p40 subunit of the Langer-Gledion syndrome region on  
RT factor eif3 has 8 exons, maps to the TRPS gene",  
RL chromosome 8q24, but is not the TRPS gene",  
[3]  
SEQUENCE FROM N.A.  
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
STRAUSBERG R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenfer C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hejblum L.,  
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carrincci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Vallat D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences",  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

CC -!- FUNCTION: Binds to the 40S ribosome and promotes the binding of
CC methionyl-tRNAi and mRNA. Associates with the p170 subunit of
CC EIF3.
CC -!- SUBUNIT: eIF-3 is composed of at least 12 different subunits.
CC -!- SIMILARITY: Contains 1 MEN (JAB/MOV34) domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U54559; AAC03465.1; -
DR EMBL: AF092576; AAC84044.1; -
DR EMBL: AF092569; AAC84044.1; JOINED.
DR EMBL: AF092570; AAC84044.1; JOINED.
DR EMBL: AF092571; AAC84044.1; JOINED.
DR EMBL: AF092572; AAC84044.1; JOINED.
DR EMBL: AF092573; AAC84044.1; JOINED.
DR EMBL: AF092574; AAC84044.1; JOINED.
DR EMBL: AF092575; AAC84044.1; JOINED.
DR EMBL: BC000386; AAC00386.1; -
DR MEROPS: M67.971; -
DR Gene: HGNC:3273; EIF3S3.
DR Reactome: O15372; EIF3S3.
DR MIM: 603912; -
DR GO: GO:0005852; C:eukaryotic translation initiation factor 3. ; TMS.
DR GO: GO:0008135; F:translation factor activity, nucleic acid b. ; TMS.
DR GO: GO:0006446; P:regulation of translational initiation; TMS.
DR InterPro: IPR000639; MOV34-1.
DR InterPro: IPR000555; MOV34_MPN_PAD1.
DR Pfam: PF01398; MOV34; 1.
DR ProDom: PD363422; MOV34_1; 1.
DR SMART: SM00232; JAB_MPN; 1.
DR Initiation factor; Protein biosynthesis.
DR CONFLICT 73 E -> K (in Ref. 2).
FT SEQUENCE 352 AA; 39930 MW; F3A6EPAOCEP587D0 CRC64;
SQ

```

```

Query Match          100.0%; Score 43; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 NLOQLMDRV 9
Db 242 NLOQLMDRV 250

```

```

RESULT 2
Q6P9U8 PRELIMINARY; PRT; 352 AA.
AC 06P9U8;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 3, subunit 3 gamma,
DE 40kDa.
DE Name=EIF3S3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Plutitary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

```

```

RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

```

SQ SEQUENCE 352 AA; 39905 MW; C06307269ADB343 CRC64;

```

```

Query Match          100.0%; Score 43; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 NLOQLMDRV 9
Db 242 NLOQLMDRV 250

```

```

RESULT 3
AAH60586 PRELIMINARY; PRT; 352 AA.
AC AAH60586;
DT 14-APR-2004 (TREMBlrel. 27, Created)
DT 14-APR-2004 (TREMBlrel. 27, Last sequence update)
DT 14-APR-2004 (TREMBlrel. 27, Last annotation update)
DE Eukaryotic translation initiation factor 3, subunit 3 gamma,
DE 40kDa.
DE Name=EIF3S3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Plutitary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Matusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```



```

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=pituitary gland;
RA Strusberg R.;
RL Submitted (OCT-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC060586; AAH60586.1; -.
KW Initiation factor.
SQ SEQUENCE 352 AA; 39905 MW; C06307269ADB343 CRC64;

Query Match
Best Local Similarity 100.0%; Score 43; DB 2; Length 352;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLOQLMDRV 9
Db 242 NLOQLMDRV 250

RESULT 4
CAG3187 PRELIMINARY; PRT; 352 AA.
ID CAG3187;
AC CAG3187;
DT 01-JUN-2004 (TREMBLrel. 27, Created)
DT 01-JUN-2004 (TREMBLrel. 27, Last annotation update)
DE 01-JUN-2004 (TREMBLrel. 27, Last annotation update)
DE EIF3S3 protein.
GN EIF3S3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RT "Cloning of human full open reading frames in Gateway(TM) system entry
RT vector (pDONR201).";
RL Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; CR456906; CAG3187.1; -.
SQ SEQUENCE 352 AA; 39930 MW; F3A6EFA0CE587D0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 43; DB 2; Length 352;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLOQLMDRV 9
Db 242 NLOQLMDRV 250

RESULT 5
Q6P381 PRELIMINARY; PRT; 335 AA.
ID Q6P381;
AC Q6P381;
DT 05-JUN-2004 (TREMBLrel. 27, Created)
DT 05-JUN-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUN-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75580.
GN Name=MGC75580;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxId=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=22388257; PubMed=12477932;
RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

```

```

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Keteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strusberg R.;
RL Submitted (DEC-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC064151; AAH64151.1; -.
DR InterPro; IPR000555; Mov34_MPN_PAD1.
DR Pfam; PF01398; Mov34; 1.
DR SMART; SM00232; JAB_MPN; 1.
KW Hypothetical protein.
SQ SEQUENCE 335 AA; 38538 MW; 1BD4D446D6561BE9 CRC64;

Query Match
Best Local Similarity 88.4%; Score 38; DB 2; Length 335;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLOQLMDRV 9
Db 225 NLOQLMDRV 233

RESULT 6
AAH64151 PRELIMINARY; PRT; 335 AA.
ID AAH64151;
AC AAH64151;
DT 25-MAR-2004 (TREMBLrel. 27, Created)
DT 25-MAR-2004 (TREMBLrel. 27, Last annotation update)
DT 25-MAR-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein MGC75580.
GN Name=MGC75580;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxId=8364;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=22388257; PubMed=12477932;
RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Keteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

```

[2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Embryo;  
 RC Klein S., Strausberg R.,  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC064151; AH64151.1;  
 KM Hypothetical protein.  
 SO SEQUENCE 335 AA; 38538 MW; 18D4D446D561B9 CRC64;  
 Query Match 88.4%; Score 38; DB 2; Length 335;  
 Best Local Similarity 88.9%; Pred. No. 24;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NIOQLMDRV 9  
 Db 225 SLQQLMDRV 233  
 RESULT 7  
 ID IF33\_MOUSE STANDARD; PRT; 352 AA.  
 AC Q91WK2;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Eukaryotic translation initiation factor 3 subunit 3 (eif-3 gamma)  
 DE (eif3 p40 subunit) (eif3h).  
 GN Name=Eif3s3;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye, and Retina;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uudin T.B., Toshilovki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Rahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Binds to the 40S ribosome and promotes the binding of  
 CC methionyl-tRNAi and mRNA. Associates with the p170 subunit of  
 CC EIF3.  
 CC -1- SUBUNIT: eif-3 is composed of at least 12 different subunits.  
 CC -1- SIMILARITY: Contains 1 MPN (UAB/Mov34) domain.  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: BC014755; AAH14755.1;  
 DR MEROPS: M67.971;  
 DR MGD: MGI:1915385; Eif3s3.

DR Interpro: IPR003639; Mov34-1.  
 DR Interpro: IPR000555; Mov34\_MPN\_PAD1.  
 DR Pfam: PF01398; Mov34; 1.  
 DR Prodom: PD363422; Mov34.1; 1.  
 DR SMART: SM00232; UAB MPN.1.  
 KM Initiation factor; Protein biosynthesis.  
 SO SEQUENCE 352 AA; 39832 MW; 96F5ABBE2F41F838 CRC64;  
 Query Match 88.4%; Score 38; DB 1; Length 352;  
 Best Local Similarity 88.9%; Pred. No. 25;  
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NIOQLMDRV 9  
 Db 242 SLQQLMDRV 250  
 RESULT 8  
 ID Q8BTK5 PRELIMINARY; PRT; 352 AA.  
 AC Q8BTK5;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-  
 DE length enriched library, clone:E430016X01 product:eukaryotic  
 DE translation initiation factor 3, subunit 3 (gamma, 40kD), full insert  
 DE sequence.  
 GN Name=Eif3s3;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.,  
 RT "High-efficiency full-length cDNA cloning."  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RA The FANTOM Consortium,  
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes."  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NOD; TISSUE=Thymus;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Atzawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Kono H., Akiyama U., Mishi K., Kitsuana T., Tashiro T., Harada A.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanuki M.,

```

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "Riken integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RN SEQUENCE FROM N.A.
RP STRAIN=NOD; TISSUE=Thymus;
RC Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuura T., Miyazaki A., Murata M., Nakamura M.,
RA Nihei K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK088436; BAC40351.1; -.
DR GO; GOI:0003743; F:translation initiation factor activity; IEA.
DR InterPro; IPR000555; MOV34_MPN_PADI.
DR Pfam; PF01398; MOV34; 1.
DR SMART; SM00232; JAB_MPN; 1.
DR Initiation factor.
KW SEQUENCE 352 AA; 39846 MW; E6F5A89E2F41E97D CRC64;
SQ
Query Match
Best Local Similarity 88.4%; Score 38; DB 2; Length 352;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 NLOQLMDRV 9
DB 242 SLQQLMDRV 250

```

```

RESULT 10
Q94BU3 PRELIMINARY; PRT; 196 AA.
AC Q94BU3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE T29M8.1/T29M8.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eucosida II; Brassicales; Brassicaceae; Arabidopsids.
OK NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.
RP Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bower L., Carninci P., Dale J.M., Gibson H.A., Goldsmith A.D.,
RA Hayashizaki Y., Ishida Y., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinzaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [12]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bower L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida Y., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY077672; AAL76150.1; -.
DR EMBL; AY077672; AAL76150.1; -.
SQ SEQUENCE 196 AA; 21478 MW; 995CA7D4BADF82AA CRC64;
QY 2 LQQLMDRV 9
DB 146 LQQLMDRV 153

```

```

Query Match
Best Local Similarity 83.7%; Score 36; DB 2; Length 196;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 LQQLMDRV 9
DB 146 LQQLMDRV 153

```

```

RESULT 11
Q8LAP2 PRELIMINARY; PRT; 311 AA.
AC Q8LAP2;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eucosida II; Brassicales; Brassicaceae; Arabidopsids.
OK NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.
RP Haas B.J., Volkovskiy N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [12]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;

```

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY087694; AAM65231.1; -  
 KM Hypothetical protein  
 SQ SEQUENCE 311 AA; 34387 MW; 1ACCD30D7260A0A0 CRC64;  
 Query Match 83.7%; Score 36; DB 2; Length 311;  
 Best Local Similarity 87.5%; Pred. No. 57;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LQIIMDRV 9  
 Db 146 LQIIMDR1 153  
 RESULT 12  
 Q9LMB2 PRELIMINARY; PRT; 311 AA.  
 AC Q9LMB2; (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE T29M8.1 protein.  
 GN Name=T29M8.1;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Liu S.X., Vaysberg M., Etgu P., Lee J.M., Lenz C., Pham P., Sakano H.,  
 RA Tortum M., Yu G., Chan A., Chung M., Goldsmith A., Liu A., Smith A.,  
 RA Altati H., Brooks S., Buehler E., Chao Q., Conn L., Conway A.B.,  
 RA Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M.,  
 RA Nguyen M., Palm C.J., Shim P., Southwick A., Davis R.W., Ecker J.R.,  
 RA Federspiel N.A., Theologis A.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC069143; AAF82225.1; -  
 DR PIR: G86324; G86324. MW; PACCASBDB69623AA0 CRC64;  
 SQ SEQUENCE 311 AA; 34341 MW; PACCASBDB69623AA0 CRC64;  
 Query Match 83.7%; Score 36; DB 2; Length 311;  
 Best Local Similarity 87.5%; Pred. No. 57;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LQIIMDRV 9  
 Db 146 LQIIMDR1 153  
 RESULT 13  
 Q9LMB3 PRELIMINARY; PRT; 342 AA.  
 AC Q9LMB3; (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE F14D16.30.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,  
 RA Kim C., Altati H., Bei Q., Chin C., Chiu J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharly N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Tortum M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altati H., Bei B., Chin C., Chiu J., Choi E.,  
 RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharly N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Tortum M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC068602; AAF79297.1; -  
 SQ SEQUENCE 342 AA; 38937 MW; 48B2B19ECF26617D CRC64;  
 Query Match 83.7%; Score 36; DB 2; Length 342;  
 Best Local Similarity 87.5%; Pred. No. 63;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 LQIIMDRV 9  
 Db 146 LQIIMDR1 153  
 RESULT 14  
 Q9NZ20 PRELIMINARY; PRT; 133 AA.  
 AC Q9NZ20; (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE HSPC180.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=Blood;  
 RC MEDLINE=2049367; PubMed=11042152;  
 RA Zhang Q.H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,  
 RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.G., Zhang J.W.,  
 RA Tao J., Huang Q.H., Zhou J., Hu G.X., Gu J., Chen S.J., Chen Z.;  
 RT "Cloning and functional analysis of cDNAs with open reading frames for  
 RT 300 previously undefined genes expressed in CD34+ hematopoietic  
 RT stem/progenitor cells.";  
 RL Genome Res. 10:1546-1560 (2000).  
 DR EMBL: AF161528; AAF29143.1; -  
 DR SWISS-2DPAGE: Q9NZ20; -  
 SQ SEQUENCE 133 AA; 15221 MW; 9C538A2DED755248 CRC64;  
 Query Match 81.4%; Score 35; DB 2; Length 133;  
 Best Local Similarity 87.5%; Pred. No. 40;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NQIIMDR 8  
 Db 23 NQIIMDR 30  
 RESULT 15  
 Q9Y221 PRELIMINARY; PRT; 180 AA.  
 AC Q9Y221; (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE HSPC031 (OK/SW-Cl.78) (CGI-37 protein) (Hypothetical protein) (OK/SW-  
 DE cl.76 protein).  
 GN Name=OK/SW-cl.78; Synonyms=OK/SW-cl.76;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=20499367; PubMed=11042152;  
 RA Zhang Q.H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,  
 RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.G., Zhang J.W.,  
 RA Tao J., Huang Q.H., Zhou J., Hu G.X., Gu J., Chen S.J., Chen Z.;  
 RT "Cloning and functional analysis of cDNAs with open reading frames for  
 RT 300 previously undefined genes expressed in CD34+ hematopoietic  
 RT stem/progenitor cells.";  
 RL Genome Res. 10:1546-1560(2000).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RA Zhang Q., Ye M., Fu G., Zhou J., Guan Z., Huang Q., Xu S., He K.,  
 RA Chen S., Shen Y., Chen Z.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RA Shichijo S., Itoh K.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20272150; PubMed=10810093;  
 RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;  
 RT "Identification of novel human genes evolutionarily conserved in  
 RT Caenorhabditis elegans by comparative proteomics.";  
 RL Genome Res. 10:703-713(2000).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Dichtchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Useth T.B., Toshiyuki S., Carrinci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Gay L.J., Hulik S.W.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Mair M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Myokai F.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF085360; AAD40195.1; -  
 DR EMBL; AF082479; BAB93504.1; -  
 DR EMBL; AF132971; AAD27746.1; -  
 DR EMBL; BC015941; AAH15941.1; -  
 DR EMBL; AB112439; BAD05056.1; -  
 DR EMBL; AB062398; BAB93485.1; -  
 DR GO:0003723; F-RNA binding; IEA.  
 DR InterPro; IPR002478; PUA.  
 DR InterPro; IPR005155; UPF0113.  
 DR Pfam; PF03657; UPF0113; 1.  
 DR SMART; SM00359; PUA; 1.  
 DR PROSITE; PSS0890; PUA; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 180 AA; 20462 MW; FCGCFB2250AA4FC9 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 54;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NLOILMDR 8  
 Db 23 NLOILMDR 30

## RESULT 16

Q9WV50 PRELIMINARY; PRT; 180 AA.

AC Q9WV50; 01-NOV-1999 (TREMBLrel. 12, Created)

DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DE Peachy.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_Taxid=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley; TISSUE=Kidney;

RA Wang Z., Tennisswood M.;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Pituitary gland;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Dichtchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Useth T.B., Toshiyuki S., Carrinci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Gay L.J., Hulik S.W.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,

RA Jones S.J., Mair M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Pituitary gland;

RA Strausberg R.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF158186; AAD42887.1; -

DR EMBL; BC005914; AAH59114.1; -

DR GO:0003723; F-RNA binding; IEA.

DR InterPro; IPR002478; PUA.

DR InterPro; IPR005155; UPF0113.

DR Pfam; PF03657; UPF0113; 1.

DR PROSITE; PSS0890; PUA; 1.

SQ SEQUENCE 180 AA; 20431 MW; B39D7B9ED85509A CRC64;

Query Match 81.4%; Score 35; DB 2; Length 180;

Best Local Similarity 87.5%; Pred. No. 54;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOILMDR 8

Db 23 NLOILMDR 30

RESULT 17  
Q9CCK8 PRELIMINARY; PRT; 180 AA.  
ID Q9CCK8;  
AC Q9CCK8;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-OCT-2004 (TREMBLrel. 28, Last annotation update)  
DE Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length enriched  
DE library, clone:320001L04 product:PBACHY homolog (Saccharomyces  
DE cerevisiae N1P7P homolog).  
DE Name=1110017C15Rik;  
GN Mus musculus (Mouse).  
OS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota; Euteleostomi; Chordata; Sclerozoa; Mollusca; Mus.  
OC Mammalia; Euteleostomi; Rodentia; Sclerozoa; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=99279253; PubMed=10349636;  
RA Carninci P., Hayashizaki Y.,  
RT "High-efficiency full-length cDNA cloning.",  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=21085660; PubMed=11217851;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.",  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RA The FANTOM Consortium,  
RT "The RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60 770 full-length cDNAs.",  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=20493374; PubMed=11042159;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.",  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RX MEDLINE=20530913; PubMed=11076861;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Kono H., Akiyama Y., Nishi K., Katsunai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.",  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Head;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Haraguchi T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,  
RA Imocant K., Ishii Y., Itoh M., Izawa M., Kato H.,  
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kuritara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Nunazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai K., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shitaki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,

RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Mix FVB/N;  
RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altshuler S.F., Zeeberg B., Buettner K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marzella K., Farmer A.A., Rubin G.M., Hong L.,  
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Tomblin S., Carninci P., Prange C.,  
RA Raha S.S., Loguercio N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Bouffard G.G.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
RA Kravinsky M.I., Skalka U., Smalls D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.",  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [8]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Mix FVB/N;  
RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;  
RA Strausberg R.,  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RX EMBL; AK014290; BAB29247.1; -;  
DR EMBL; BC003972; AA003972.1; -;  
DR GSI; MGI:1913414; 1110017C15Rik.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR InterPro; IPR002478; PUF.  
DR InterPro; IPR005155; UPP0113.  
DR Pfam; PF03657; UPP0113; 1.  
DR SMART; SM00359; PUA; 1.  
DR PROSITE; PS50890; PUA; 1.  
SQ SEQUENCE 180 AA; 20451 MW; 4F1256165370F2AF CRC64;  
Query Match 81.4%; Score 35; DB 2; Length 180;  
Best Local Similarity 87.5%; Pred. No. 54;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
RESULT 18  
ID Q9DIB4 PRELIMINARY; PRT; 180 AA.  
AC Q9DIB4;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched  
DE library, clone:1110017C15Rik product:PBACHY homolog.  
DE Name=1110017C15Rik;  
GN Mus musculus (Mouse).  
OS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota; Euteleostomi; Chordata; Sclerozoa; Mollusca; Mus.  
OC Mammalia; Euteleostomi; Rodentia; Sclerozoa; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=99279253; PubMed=10349636;  
RA Carninci P., Hayashizaki Y.,  
RT "High-efficiency full-length cDNA cloning.",

RL Mech. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=whole body;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=whole body;  
 RA The FANTOM Consortium,  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=whole body;  
 RX MEDLINE=20499374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes."  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=whole body;  
 RX MEDLINE=20530913; PubMed=11076861;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,  
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
 RA Yamamoto S., Matsumoto H., Sakaguchi S., Ikegami T., Kasahigagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer."  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=whole body;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamuro T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.,  
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AK003739; BAB22972.1;  
 DR MGD; MGI:1913414; 1110017C1SR1k.  
 DR GO; GO:0003723; F:RNA binding; IEA.  
 DR InterPro; IPR002478; PUA.  
 DR InterPro; IPR005155; UPF0113.  
 DR Pfam; PF03657; UPF0113; 1.  
 DR SMART; SM00359; PUA; 1.  
 DR PROSITE; PS50890; PUA; 1.  
 SQ SEQUENCE 180 AA; 20392 MW; 4C9FE1E18D209885 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 54;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOQLMDR 8  
 |||||  
 DB 23 NLOQLVDR 30

RESULT 19  
 ID AAH59114 PRELIMINARY; PRT; 180 AA.  
 AC AAH59114;  
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Peachy protein.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=pituitary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Hellon E., Kettleman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butlerfield Y.S.,  
 RA Krzywinski M.I., Skalske U., Smaltis D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Mair M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=pituitary gland;  
 RA Strausberg R.,  
 RT Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC059114; AAH59114.1;  
 SQ SEQUENCE 180 AA; 20431 MW; B39D7B9ED8E5509A CRC64;

Query Match 81.4%; Score 35; DB 2; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 54;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLOQLMDR 8  
 |||||  
 DB 23 NLOQLVDR 30

RESULT 20  
 ID BAD05056 PRELIMINARY; PRT; 180 AA.  
 AC BAD05056;  
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Myokai F.,  
 RT "Mechanical stress-induced gene 26-1, 26-2."  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB112439; BAD05056.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 180 AA; 20462 MW; FC6CFB2250AAA4FC9 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 180;  
 Best Local Similarity 87.5%; Pred. No. 54;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLOQLMDR 8  
 DB 23 NLOQLMDR 30

## RESULT 21

OY 07T21 PRELIMINARY; PRT; 1742 AA.

AC 07T21; 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Tsc2 protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Colling E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences."  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC052449; AAH52449.1;  
 DR EMBL; BC060701; AAH60701.1;  
 DR GO; GO:0046626; P:regulation of insulin receptor signaling pa. .; IMP.  
 DR InterPro; IPR000331; Rap GAP.  
 DR InterPro; IPR003913; Tuberin.  
 DR Pfam; PF02145; Rap-GAP; 1.  
 DR Pfam; PF03542; Tuberin; 1.  
 DR PROSITE; PSS0085; RAPGAP; 1.  
 DR PROSITE; PSS0085; RAPGAP; 1.  
 SO SEQUENCE 1742 AA; 194097 MW; 0515487278578EF2 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 1742;  
 Best Local Similarity 87.5%; Pred. No. 4.9e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLOQLMDR 8  
 DB 444 NLOQLMDR 451

RESULT 22  
 ID AAH60701 PRELIMINARY; PRT; 1742 AA.

AC AAH60701; 02-MAR-2004 (TREMBLrel. 27, Created)  
 DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)  
 DE Tsc2 protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Colling E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences."  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC060701; AAH60701.1;  
 DR EMBL; BC060701; AAH60701.1;  
 SO SEQUENCE 1742 AA; 194097 MW; 0515487278578EF2 CRC64;

Query Match 81.4%; Score 35; DB 2; Length 1742;  
 Best Local Similarity 87.5%; Pred. No. 4.9e+02;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NLOQLMDR 8  
 DB 444 NLOQLMDR 451

## RESULT 23

ID TSC2\_RAT STANDARD; PRT; 1809 AA.

AC P49616; 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Tuberin (Tuberous sclerosis 2 homolog protein).  
 GN Name=Tsc2;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1, 2, 3 AND 4).  
 RC STRAIN=Miscar; TISSUE=Kidney;  
 RX MEDLINE=96063895; PubMed=8519695;  
 RA Xiao G.-H., Jin F., Yeung R.S.;



RT "Identification of tuberous sclerosis 2 messenger RNA splice variants  
RT that are conserved and differentially expressed in rat and human  
RT tissues.";  
RN Cell Growth Differ. 6:1185-1191(1995).  
[2]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=Long Evans; TISSUE=Brain, and Kidney;  
RX MEDLINE=9380273; PubMed=7651821;  
RA Kobayashi T., Nishizawa M., Hirayama Y., Kobayashi E., Hino O.;  
RT "CDNA structure, alternative splicing and exon-intron organization of  
RT the predisposing tuberous sclerosis (Tsc2) gene of the Eker rat  
RT model";  
RL Nucleic Acids Res. 23:2608-2613(1995).  
[3]  
RN TUBERIN-RAB5 BINDING.  
RP MEDLINE=97197768; PubMed=9045618;  
RX Xiao G.-H., Shoarinejad F., Jin F., Golemis E.A., Yeung R.S.;  
RT "The tuberous sclerosis 2 gene product, tuberin, functions as a Rabs  
RT GTPase activating protein (GAP) in modulating endocytosis.";  
RL J. Biol. Chem. 272:6097-6100(1997).  
[4]  
RN MUTAGENESIS OF ASN-314 AND LEU-713.  
RX MEDLINE=99151532; PubMed=10029074;  
RA Sarake N., Kobayashi T., Kobayashi E., Izumi K., Hino O.;  
RT "Isolation and characterization of a rat homologue of the human  
RT renal carcinoma 1 gene (Tsc1) and analysis of its mutations in rat  
RT renal carcinoma.";  
RL Cancer Res. 59:849-855(1999).  
-1- FUNCTION: Implicated as a tumor suppressor. May have a function in  
CC vesicular transport, but may also play a role in the regulation of  
CC cell growth arrest and in the regulation of transcription mediated  
CC by steroid receptors. Interaction between hamartin and tuberin may  
CC facilitate vesicular docking. Specifically stimulates the  
CC intrinsic GTPase activity of the ras-related protein RAP1A and  
CC RAS. Suggesting a possible mechanism for its role in regulating  
CC cellular growth.  
-1- SUBUNIT: Interacts with hamartin. May also interact with the  
CC adaptor molecule rabapin 5. The final complex contains tuberin  
CC and rabapin 5 linked to RAB5 (Probable).  
-1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=4;  
CC Comment=Additional isoforms seem to exist;  
CC Name=1;  
CC IsoId=P49816-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=P49816-2; Sequence=VSP\_004481;  
CC Name=3;  
CC IsoId=P49816-3; Sequence=VSP\_004482;  
CC Name=4;  
CC IsoId=P49816-4; Sequence=VSP\_004481, VSP\_004482;  
-1- TISSUE SPECIFICITY: CNS, uterus, heart, skeletal muscle, kidney  
CC and spleen.  
-1- DISEASE: A germline insertion in Tsc2 is the cause of the Eker rat  
CC model of inherited cancer susceptibility. Gives rise to a spectrum  
CC of epithelial and nonepithelial neoplasms.  
-1- SIMILARITY: Contains 1 Rap-GAP domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
-----  
CC EMBL; U24150; AAC52289.1; -;  
DR EMBL; D50413; BAA08914.1; -;  
DR PIR; S57329; S57329.  
DR RGD; 3908; Tsc2.  
DR InterPro; IPR008938; ARM.  
DR InterPro; IPR003913; Rap\_GAP.  
DR InterPro; IPR003913; Tuberin.  
DR Pfam; PF02145; Rap\_GAP; 1.

DR Pfam; PF03542; Tuberin; 1.  
DR PRINTS; PR01431; TUBERIN.  
DR PROSITE; PS50085; RAPGAP; 1.  
KW Alternative splicing; Anti-oncogene; GTPase activation.  
FT DOMAIN 1533 1760  
FT VARSPPLIC 947 989  
FT VARSPPLIC 1272 1294  
FT VARSPPLIC 1272 1294  
FT MUTAGEN 314 314  
FT MUTAGEN 713 713  
FT MUTAGEN 932 932  
FT CONFLICT 1514 1514  
FT CONFLICT 1730 1730  
SQ SEQUENCE 1809 AA; 201276 MW; 6190BEP45272664 CRC64;  
Query Match 81.4%; Score 35; DB 1; Length 1809;  
Best Local Similarity 87.5%; Pred. No. 5, 1e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NQGLMDR 8  
Db 444 NQGLMDR 451  
RESUL 24  
TSC2\_MOUSE STANDARD; PRT; 1814 AA.  
ID TSC2\_MOUSE PRT; 1814 AA.  
AC Q61037; P97723; P97724; P97727; Q61007; Q61008; Q9WUF6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Tuberin (Tuberous sclerosis 2 homolog protein).  
GN Name=Tsc2;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS A; B; C; D; E; F AND G).  
RC TISSUE=Heart;  
RX MEDLINE=96258425; PubMed=8777431;  
RA Kim K.K., Patlak L., Wang H., Field L.J.;  
RT "Cloning, developmental expression, and evidence for alternative  
RT splicing of the murine tuberous sclerosis (TSC2) gene product.";  
RL Cell. Mol. Biol. Res. 41:515-526(1995).  
[2]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RX MEDLINE=96430093; PubMed=8833243;  
RA Olsson P.G., Schofield J.N., Edwards Y.H., Friesch A.M.;  
RT "Expression and differential splicing of the mouse TSC2 homolog.";  
RL Mamm. Genome 7:212-215(1996).  
[3]  
RP SEQUENCE OF 1-199 FROM N.A.  
RC STRAIN=BA1B/C; TISSUE=Leukocyte;  
RX MEDLINE=98417643; PubMed=9743625;  
RA Sarter A.H., Ikeda S., Nakano H., Terato H., Ide H., Imai K.,  
RA Akiyama K., Tsubui K., Bo Z., Kubo K., Yamamoto K., Yasui A.,  
RA Yoshiida M.C., Seki S.;  
RT "Cloning and characterization of a mouse homologue (mTsc1) of  
RT Escherichia coli endonuclease III.";  
RL J. Mol. Biol. 262:761-774(1998).  
[4]  
RP SEQUENCE OF 119-1805 FROM N.A.  
RX MEDLINE=20051947; PubMed=10584558;  
RA Kleymanova B.V., Declue J.E., Walker C.L.;  
RT "Genetic variants of the tuberous sclerosis 2 tumour suppressor gene  
RT in mouse t haplotypes".  
RL Genet. Res. 74:139-144(1999).  
-1- FUNCTION: Implicated as a tumor suppressor. May have a function in  
CC vesicular transport, but may also play a role in the regulation of



RA Chen S.L., Xu M.Y.;  
 RT "Analysis of expressed genes in red sea bream (*Chrysophrys major*).";  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY10743; AAP20218.1; -  
 KW Initiation factor.  
 FT NON TER 149  
 SQ SEQUENCE 149 AA; 17246 MW; 26DA91A07DF0CB79 CRC64;  
 Query Match  
 Best Local Similarity 79.1%; Score 34; DB 2; Length 149;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 NLOQLMDR 8  
 Db 142 SIQQLMDR 149  
 RESULT 27  
 Q31388 PRELIMINARY; PRT; 285 AA.  
 AC Q31388;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE MHC class I protein precursor (Fragment).  
 OS *Cyprinus carpio* (Common carp).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Cyprinus.  
 OX NCBI\_Taxid=7962;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DOR70;  
 RX MEDLINE=9621755; PubMed=8613142;  
 RA van Erp S.H.; Dixon B.; Figueroa F.; Egberts E.; Sret R.J.;  
 RT "Identification and characterization of a novel class I gene in carp  
 (Cyprinus carpio L.).";  
 RL Immunogenetics 44:49-61(1996).  
 DR EMBL; X91022; CAA62498.1; -  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0006955; P:immune response; IEA.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003597; IG-cl.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR InterPro; IPR01039; MHC I.  
 DR Pfam; PF07654; CI-sect; 1.  
 DR Pfam; PF00129; MHC I; 1.  
 DR PRINTS; PRO1638; MHCCLASSI.  
 DR SMART; SM00407; IGC1; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
 KW SIGNAL.  
 FT SIGNAL. 1  
 FT NON TER 285  
 SQ SEQUENCE 285 AA; 32801 MW; 74633B20688B9B82 CRC64;  
 Query Match  
 Best Local Similarity 79.1%; Score 34; DB 2; Length 285;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 NLOQLMDR 8  
 Db 91 NIQVLMR 98  
 RESULT 28  
 Q3MF8 PRELIMINARY; PRT; 560 AA.  
 AC Q3MF8;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Phosphofructokinase, pyrophosphate-dependent.

GN OrderedlocusNames=TDB1550;  
 OS Treponema denticola.  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_Taxid=158;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35405 / DSM 14222;  
 RX PubMed=15064399; DOI=10.1073/pnas.0307639101;  
 RA Sehnadri R., Myers G.S.A., Tetteijn H., Eisen J.A., Heidelberg J.F.,  
 RA Dodson R.J., Davidson T.M., Debby R.T., Fouts D.E., Haft D.H.,  
 RA Selengut J., Ren O., Brinkac L.M., Madupu R., Kolonay J.F.,  
 RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,  
 RA Gebreyeorgis E., Geer K., Tsagaye G., Malek J.A., Ayodeji B.,  
 RA Shatsman S., McLeod M.P., Smajd D., Howell J.K., Pal S., Amin A.,  
 RA Vashlish P., McNeill T.Z., Xiang O., Sodergren E., Baca E.,  
 RA Weinstein G.M., Norris S.J., Fraser C.M., Paulsen I.T.;  
 RT "Comparison of the genome of the oral pathogen *Treponema denticola*  
 with other spirochete genomes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).  
 DR EMBL; AE017251; AAS12067.1; -  
 DR TIGR; TDB1550; -  
 DR GO; GO:0016301; F:kinase activity; IEA.  
 DR InterPro; IPR01183; Pfam\_PPI\_PFK.  
 DR InterPro; IPR00023; Pfructkinase.  
 DR Pfam; PF00365; PFK; 1.  
 DR PIRSF; PIRSF005677; Pfam\_PPI\_PFK; 1.  
 DR PRINTS; PR00476; PHFRCTKINASE.  
 DR ProDom; PD000707; Pfructkinase; 1.  
 KW Complete proteome; kinase.  
 SQ SEQUENCE 560 AA; 63141 MW; D5D9695D26343B17 CRC64;  
 Query Match  
 Best Local Similarity 79.1%; Score 34; DB 2; Length 560;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 NLOQLMDR 8  
 Db 371 SIQQLMDR 378  
 RESULT 29  
 AAS12067 PRELIMINARY; PRT; 560 AA.  
 AC AAS12067;  
 ID AAS12067;  
 DT 02-MAR-2004 (TREMBLrel. 27, Created)  
 DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)  
 DT 04-MAY-2004 (TREMBLrel. 27, Last annotation update)  
 DE Phosphofructokinase, pyrophosphate-dependent.  
 GN TDB1550.  
 OS Treponema denticola.  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_Taxid=158;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35405 / DSM 14222;  
 RX PubMed=15064399;  
 RA Sehnadri R., Myers G.S.A., Tetteijn H., Eisen J.A., Heidelberg J.F.,  
 RA Dodson R.J., Davidson T.M., Debby R.T., Fouts D.E., Haft D.H.,  
 RA Selengut J., Ren O., Brinkac L.M., Madupu R., Kolonay J.F.,  
 RA Daugherty S.C., Shetty J., Shvartsbeyn A., Gebreyeorgis E., Geer K.,  
 RA Tsagaye G., Malek J.A., Ayodeji B., Shatsman S., McLeod M.P.,  
 RA Smajd D., Howell J.K., Pal S., Amin A., Vashlish P., McNeill T.Z.,  
 RA Xiang O., Sodergren E., Baca E., Weinstein G.M., Norris S.J.,  
 RA Fraser C.M., Paulsen I.T.;  
 RT "Comparison of the genome of the oral pathogen *Treponema denticola*  
 with other spirochete genomes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).  
 DR EMBL; AE017251; AAS12067.1; -  
 DR TIGR; TDB1550; -  
 KW kinase.  
 SQ SEQUENCE 560 AA; 63141 MW; D5D9695D26343B17 CRC64;  
 Query Match  
 Best Local Similarity 79.1%; Score 34; DB 2; Length 560;

Best Local Similarity 87.5%; Pred. No. 2.6e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLOQLMDR 8  
Db 371 NLOQLMDR 378

## RESULT 30

RPOC ANTFO STANDARD; PRT; 681 AA.  
AC Q85C16;  
DT 29-MAR-2004 (Rel. 43, Created)  
DT 29-MAR-2004 (Rel. 43, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (PPT) (plastid-  
encoded RNA polymerase beta subunit) (RNA polymerase beta subunit).  
GN Name=rpoC1;  
OS Anthoceros formosae (Hornwort).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Anthocerotophyta;  
OC Anthocerotales; Anthocerotaceae; Anthoceros.  
OX NCBI\_TaxID=48387;  
[1]  
RP SEQUENCE FROM N.A., AND RNA EDITING.  
RP TISSUE=Thallus;  
RX MEDLINE=22415709; PubMed=12527781;  
RA Kushta M., Kaneko A., Yamamoto Y., Takeya Y., Matsumoto T.,  
RA Yoshinaga K.;  
RT "The complete nucleotide sequence of the hornwort (Anthoceros  
formosae) chloroplast genome: insight into the earliest land plants.";  
RL Nucleic Acids Res. 31:716-721(2003).  
RN [2]  
RP SEQUENCE FROM N.A., AND RNA EDITING.  
RP TISSUE=Thallus;  
RX Kushta M., Yamamoto Y., Fujikawa T., Matsumoto T., Yoshinaga K.;  
RT "RNA editing in hornwort chloroplasts makes more than half the genes  
functional." Res. 31:2417-2423(2003).  
RL Nucleic Acids Res. 31:2417-2423(2003).  
CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription  
of DNA into RNA using the four ribonucleoside triphosphates as  
substrates.  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
[RNA] (N).  
CC -1- SUBUNIT: In plastids the minimal PEP RNA polymerase catalytic core  
is composed of four subunits: alpha, beta, beta', and beta''. When  
a (nuclear-encoded) sigma factor is associated with the core the  
holoenzyme is formed, which can initiate transcription (By  
similarity).  
CC -1- SUBCELLULAR LOCATION: Chloroplast.  
CC -1- RNA EDITING: Modified positions=43, 56, 87, 123, 238, 240, 270,  
273, 290, 296, 335, 341, 344, 355, 368, 375, 384, 390, 398, 406,  
435, 457, 466, 477, 481, 486, 494, 512, 523, 548, 653, 654;  
CC Note=The nonsense codons at positions 240, 344, 375, 494 and 548  
are modified to sense codons.  
CC -1- SIMILARITY: Belongs to the RNA polymerase beta' chain family.  
CC RPOC1 subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AB086179; BACS5327.1; -;  
CC EMBL: AB087419; BACS5418.1; -;  
CC HSP: O9KM06; IHOM.  
CC HAMAP: MF\_01323; -; 1.  
CC InterPro: IPR006592; RNA\_pol\_A.  
CC InterPro: IPR000722; RNA\_pol\_A.

DR InterPro: IPR007080; RNA\_pol\_Rpb1\_1.  
DR InterPro: IPR007066; RNA\_pol\_Rpb1\_3.  
DR Pfam: PF04997; RNA\_pol\_Rpb1\_1; 1.  
DR Pfam: PF00623; RNA\_pol\_Rpb1\_2; 1.  
DR Pfam: PF04983; RNA\_pol\_Rpb1\_3; 1.  
DR SMART: SM00663; RPOA\_N; 1.  
KW Chloroplast; DNA-directed RNA polymerase; RNA editing; Transcription;  
KW Transferase.  
SQ SEQUENCE 681 AA; 78871 MW; 164735C16338CDB CRC64;

Query Match 79.1%; Score 34; DB 1; Length 681;  
Best Local Similarity 75.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLOQLMDR 8  
Db 207 NLOQLMDR 214

## RESULT 31

06B211 PRELIMINARY; PRT; 808 AA.  
ID 06B211  
AC 06B211;  
DT 01-OCT-2004 (TrEMBLrel. 28, Created)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Chromosome A of strain CBS767 of Debaryomyces hansenii.  
GN ORFNames=DEHA0A054569;  
OS Debaryomyces hansenii (Yeast) (Torulasporea hansenii).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.  
OX NCBI\_TaxID=4959;  
RN [1]  
RP SEQUENCE FROM N.A.  
RP STRAIN=CBS767;  
RG GENOLABURS;  
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,  
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,  
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,  
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
RA Boissarie A., Boyer J., Cartolico L., Confanioleri F., de Daruvar A.,  
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
RA Hantaye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma U., Muller H.,  
RA Nicoud J.M., Nikoleki M., Ozias S., Ozier-Kalogeropoulos O.,  
RA Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
RA Swennene D., Tekala F., Mesolowski-Jouvet M., Westhof R., Wirth B.,  
RA Zentou-Weyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,  
RA Wincker P., Souciet J.L.;  
RT "Genome evolution in yeasts.";  
RL Nature 430:35-44(2004).  
RN [2]  
RP SEQUENCE FROM N.A.  
RP STRAIN=CBS767;  
RC Genoscope;  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL: CR382133; CAG84513.1; -;  
SQ SEQUENCE 808 AA; 91224 MW; C7517D3AD2CABF67 CRC64;

Query Match 79.1%; Score 34; DB 2; Length 808;  
Best Local Similarity 77.8%; Pred. No. 3.7e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NLOQLMDR 9  
Db 792 NLOQLMDR 800

## RESULT 32

Q7U58 PRELIMINARY; PRT; 964 AA.  
ID Q7U58  
AC Q7U58;

DT 01-OCT-2003 (TREMBLrel. 25, Created)  
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Similar to response regulator aspartate phosphatase rapC.  
GN OrderedlocusNames=RBI1084;  
OS Rhodospirillum rubrum.  
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
OC Planctomycetaceae; Pirellula.  
NCBI\_TaxID=117;  
RN NCB1  
RP SEQUENCE FROM N.A.  
RC STRAIN=1;  
RX MEDLINE=22735913; PubMed=12835416;  
RA Glockner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,  
RA Schlesner H., Amann R., Reinhardt R.;  
RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
RT strain 1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
DR EMBL; BX294152; CAD77153.1; -;  
DR InterPro; IPR000345; CYC\_heme\_BS.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_5.  
KW Complete proteome.  
SQ SEQUENCE 964 AA; 109575 MW; 95158629B29AFPL CRC64;

Query Match 79.1%; Score 34; DB 2; Length 964;  
Best Local Similarity 77.8%; Pred. No. 4.4e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 1 NLOQLMDRV 9  
Db 879 NLOQLMDRV 887

## RESULT 33

ID 018415 PRELIMINARY; PRT; 1116 AA.

AC 018415; DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Nuclear protein SA.

GN Name=SA;

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI\_TaxID=7227;

RN NCB1

RP SEQUENCE FROM N.A.

RC STRAIN=Canton S.;

RA Valdeolmillos A.M., Villares R., Buena J.M., Gonzalez-Crespo S.,

RA Martinez A., Barbero J.L.;

RT "Molecular Cloning and Expression of Stromalin Protein from Drosophila

RT melanogaster: Homologous to Mammalian Stromalin Family of Nuclear

RT Proteins";

RL DNA Cell Biol. 8:699-706(1998).

DR EMBL; Y14277; CA74654.1; -.

DR PIR; T13854; T13854.

DR FLYBase; FBgn020616; SA.

DR GO; GO:0005634; C:nucleus; IEA.

DR InterPro; IPR008938; ARM.

DR InterPro; IPR005032; STRG.

DR Pfam; PF03365; STRG; 1.

KW Nuclear protein.

SQ SEQUENCE 1116 AA; 128793 MW; 1802EB9P3C7FB295 CRC64;

Query Match 79.1%; Score 34; DB 2; Length 1116;

Best Local Similarity 66.7%; Pred. No. 5.1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Cy 1 NLOQLMDRV 9  
Db 879 NLOQLMDRV 887

Db 630 NLOQLMDRV 638

## RESULT 34

ID 09VM62 PRELIMINARY; PRT; 1127 AA.

AC 09VM62; Q95RF8; DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)

DE CG3423-PA (LD34181P).

GN Name=SA;

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI\_TaxID=7227;

RN NCB1

RP SEQUENCE FROM N.A.

RC MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,

RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,

RA Burtis K.C., Busam D.A., Butler H., Cadden E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,

RA Fostler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glaser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heilmann T.J., Hernandez J.R., Houck J.,

RA Hosain D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,

RA Kimmel B., Kodira C.D., Kraft C., Kravitz S., Kilp D., Lai Z.,

RA Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclab J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D., Scheefer F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier B., Spreading A.C., Stapleton M., Strong R., Sun E.,

RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., Woodger, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,

RA Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RT "The genome sequence of Drosophila melanogaster.";

RT Science 287:2185-2195(2000).

RL Science

RN NCB1

RP SEQUENCE FROM N.A.

RC MEDLINE=22426070; PubMed=12537573;

RA Celinker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,

RA Patel S., Adams M., Champe M., Dugan S.P., Frisoe E., Hodgson A.,

RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,

RA Paclab J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,

RA Svirekas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,

RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,

RT "Finishing a whole-genome shotgun: release 3 of the Drosophila

RT melanogaster euchromatic genome sequence.";

RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

RL Genome Biol.

RN NCB1

RP SEQUENCE FROM N.A.

RC MEDLINE=22426070; PubMed=12537573;

RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
 RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celinker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genomic perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Miura S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bergman B.P.,  
 RA Betencourt B.R., Celinker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX FLYBASE;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX FLYBASE;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Friese E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuno C., Lewis S.E., Rubin G.M., Celinker S.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AEO03615; AAF52463.2;  
 DR EMBL; AY061410; AAL28958.1;  
 DR FLYBASE; FBgn0020616; SA.  
 DR InterPro; IPR008938; ARM.  
 DR InterPro; IPR005032; STAG.  
 DR Pfam; PF03365; STAG; 1.  
 SQ SEQUENCE 1127 AA; 130115 MW; 8DCC0C6934228712 CRC64;

Query Match 79.1%; Score 34; DB 2; Length 1127;  
 Best Local Similarity 66.7%; Pred. No. 5.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NLOLMDRV 9  
 Db 641 NLOALMDRI 649

RESULT 35  
 AAF52463 PRELIMINARY; PRT; 1127 AA.  
 AC AAF52463;  
 DT 01-APR-2004 (TrEMBLrel. 27, Created)  
 DT 01-APR-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 01-APR-2004 (TrEMBLrel. 27, Last annotation update)  
 DE CG3423-PA.  
 GN SA OR CG3423.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; *Drosophila*.  
 OX NCBITaxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Norton J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA April J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ileguam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kienison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris D., Mostrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Skupski M.P., Smith T.,  
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weinstein J.,  
 RA Williams S.M., Woodruff M., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celinker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Friese E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RA Swinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*  
 RT melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Miura S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bergman B.P.,  
 RA Betencourt B.R., Celinker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
 RA Patel S., Friese E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celinker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
 RT a genomic perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RX FLYBASE;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

RN [6]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (MAR-2004) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AE003615; AAF52463.2; -  
DR FlyBase; FBgn020616; SA.  
SQ SEQUENCE 1127 AA; 130115 MW; 8DCC0C6934228712 CRC64;  
Query Match  
Best Local Similarity 79.1%; Score 34; DB 2; Length 1127;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NLOQLMDRV 9  
DB 641 NLOQLMDRI 649  
RESULT 36  
ID Q9M8R8 PRELIMINARY; PRT; 257 AA.  
AC Q9M8R8;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Orf 21.  
OS Staphylococcus aureus prophage phiPV83.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
OX NCBI\_TaxID=129009;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=P83;  
RX MEDLINE=98067870; PubMed=9404084;  
RA Kaneo J., Muramoto K., Kamio Y.;  
RT "Gene of LukF-PV-like component of Panton-Valentine leukocidin in  
RT Staphylococcus aureus P83 is linked with lukM";  
RL Biosci. Biotechnol. Biochem. 61:1960-1962(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=P83;  
RA Zou D., Kaneo J., Narita S., Kamio Y.;  
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AB044554; BA97828.1; -  
SQ SEQUENCE 257 AA; 30023 MW; AC948888FDA94752 CRC64;  
Query Match  
Best Local Similarity 76.7%; Score 33; DB 2; Length 257;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NLOQLMDRV 9  
DB 151 NLOQLMDRI 159  
RESULT 37  
ID Q6F7M9 PRELIMINARY; PRT; 345 AA.  
AC Q6F7M9;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames=ACIAD3256;  
OS Acinetobacter sp. (strain ADP1).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Moraxellaceae; Acinetobacter.  
OX NCBI\_TaxID=62977;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Babbe V., Vallenet D., Kreimeyer A., Ozias S.,  
RA Labarre L., Criveller S., Robert C., Duprat S., Wincker P.,  
RA Ornston L.N., Weissbach J., Mariere P., Cohen G.N., Medigue C.;  
RT "Unique features revealed by the genome sequence of Acinetobacter sp.  
RT ADP1, a versatile and naturally transformation competent bacterium.";

RL Nucleic Acids Res. 0:0-0(2004).  
DR EMBL; CR543861; CAG69936.1; -  
DR InterPro; IPR000157; TIR.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 345 AA; 40855 MW; AA502BABC732F85C CRC64;  
Query Match  
Best Local Similarity 76.7%; Score 33; DB 2; Length 345;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NLOQLMDRV 9  
DB 166 NLOQLMDRI 174  
RESULT 38  
ID Q8TWC3 PRELIMINARY; PRT; 509 AA.  
AC Q8TWC3;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Uncharacterized protein specific for M.kandleri, MK-8 family.  
GN OrderedLocustNames=MK1112;  
OS Methanopyrus kandleri.  
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
OC Methanopyrus.  
OX NCBI\_TaxID=2320;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AV19 / DSM 6324 / JCM 9639;  
RX MEDLINE=21927647; PubMed=11930014;  
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,  
RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
RA Natalie D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Sreter K.O.,  
RA Mal'kh A.G., Koonin E.V., Kozayvkin S.A.;  
RT "The complete genome of hyperthermophilic Methanopyrus kandleri AV19  
RT and monophyly of archaeal methanogens";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
DR EMBL; AE010400; AAM02325.1; -  
KW Complete proteome.  
SQ SEQUENCE 509 AA; 58194 MW; F7D8A21F7A734AED CRC64;  
Query Match  
Best Local Similarity 76.7%; Score 33; DB 2; Length 509;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 NLOQLMDRV 9  
DB 112 NLOQLMDRI 120  
RESULT 39  
ID Q9XGCL PRELIMINARY; PRT; 874 AA.  
AC Q9XGCL;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Starch synthase, isoform V (BC 2.4.1.21).  
OS Vigna unguiculata (Cowpea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC Euphorbiales; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.  
OX NCBI\_TaxID=3917;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bhullar S.S., Willmitzer L., Kossmann U.;  
RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AJ006752; CAB40375.1; -  
DR GO; GO:0009011; F:starch synthase activity; IEA.  
DR GO; GO:0016757; P:transferase activity, transferring glycosyl. . .  
DR GO; GO:0009058; P:biosynthesis; IEA.

DR InterPro: IPR001296; Glyco trans 1.  
 DR Pfam: PF00534; Glycosyltransferase\_1; 1.  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 874 AA; 98741 MW; 8561D742868C8399 CRC64;

Query Match 76.7%; Score 33; DB 2; Length 874;  
 Best Local Similarity 75.0%; Pred. No. 6.5e+02;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLOLLMDR 8  
 |||||:  
 Db 195 NLOLLMDK 202

## RESULT 40

O88G77 PRELIMINARY; PRT; 35 AA.  
 AC O88G77;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN Ordered locus names=PP3848;  
 OS Pseudomonas putida (strain KT2440).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OK NCBI\_TaxID=160488;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22423060; PubMed=12534463;  
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
 Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
 Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,  
 Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,  
 Hance I., Chris Lee P., Holtzapfe E.K., Scanlan D., Tran K.,  
 Moazzez A., Utecherback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,  
 Medler H., Lauber J., Stjepandic D., Hobeisel J., Straetz M., Heim S.,  
 Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tummeler B.,  
 Frazer C.M.;  
 RT "Complete genome sequence and comparative analysis of the  
 metabolically versatile Pseudomonas putida KT2440.";  
 RL Environ. Microbiol. 4:799-808(2002).  
 DR EMBL; AE016788; AAN69442.1; -.  
 DR TIGR; PP3848; -.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 35 AA; 3993 MW; 607A9B5FCD40A051 CRC64;

Query Match 74.4%; Score 32; DB 2; Length 35;  
 Best Local Similarity 66.7%; Pred. No. 46;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLOLLMDRV 9  
 |||||:  
 Db 4 NLSLLDKV 12

Search completed: January 12, 2005, 20:14:16  
 Job time : 88 secs